

Certificate of Mailing

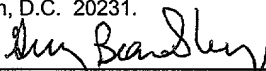
Date of Deposit November 16, 2001

Label Number: EL509047264US

I hereby certify under 37 C.F.R. § 1.10 that this correspondence is being deposited with the United States Postal Service as "Express Mail Post Office to Addressee" with sufficient postage on the date indicated above and is addressed to BOX PATENT APPLICATION, Assistant Commissioner for Patents, Washington, D.C. 20231.

Guy Beardsley

Printed name of person mailing correspondence


Signature of person mailing correspondence

APPLICATION
FOR
UNITED STATES LETTERS PATENT

APPLICANTS : Rainer Haas, Harold Kleanthous, Jean-Francois Tomb,
Charles Miller, Amal Al-Garawi, Stefan Odenbreit, and
Thomas Meyer

TITLE : Helicobacter Polypeptides and Corresponding
Polynucleotide Molecules

HELICOBACTER POLYPEPTIDES AND CORRESPONDING
POLYNUCLEOTIDE MOLECULES

5 The invention relates to *Helicobacter* antigens and corresponding polynucleotide molecules that can be used in methods to prevent or treat *Helicobacter* infection in mammals, such as humans.

Background of the Invention

10 *Helicobacter* is a genus of spiral, gram-negative bacteria that colonize the gastrointestinal tracts of mammals. Several species colonize the stomach, most notably *H. pylori*, *H. heilmanii*, *H. felis*, and *H. mustelae*. Although *H. pylori* is the species most commonly associated with human infection, *H. heilmanii* and *H. felis* have also been isolated from humans, but at lower frequencies than *H. pylori*. *Helicobacter* infects over 50% of adult populations in developed countries and
15 nearly 100% in developing countries and some Pacific rim countries, making it one of the most prevalent infections worldwide.

Helicobacter is routinely recovered from gastric biopsies of humans with histological evidence of gastritis and peptic ulceration. Indeed, *H. pylori* is now recognized as an important pathogen of humans, in that the chronic gastritis it
20 causes is a risk factor for the development of peptic ulcer diseases and gastric carcinoma. It is thus highly desirable to develop safe and effective vaccines for preventing and treating *Helicobacter* infection.

 A number of *Helicobacter* antigens have been characterized or isolated. These include urease, which is composed of two structural subunits of

approximately 30 and 67 kDa (Hu *et al.*, Infect. Immun. 58:992, 1990; Dunn *et al.*, J. Biol. Chem. 265:9464, 1990; Evans *et al.*, Microbial Pathogenesis 10:15, 1991; Labigne *et al.*, J. Bact., 173:1920, 1991); the 87 kDa vacuolar cytotoxin (VacA) (Cover *et al.*, J. Biol. Chem. 267:10570, 1992; Phadnis *et al.*, Infect. Immun. 62:1557, 1994; WO 93/18150); a 128 kDa immunodominant antigen associated with the cytotoxin (CagA, also called TagA; WO 93/18150; U.S. Patent No. 5,403,924); 13 and 58 kDa heat shock proteins HspA and HspB (Suerbaum *et al.*, Mol. Microbiol. 14:959, 1994; WO 93/18150); a 54 kDa catalase (Hazell *et al.*, J. Gen. Microbiol. 137:57, 1991); a 15 kDa histidine-rich protein (Hpn) (Gilbert *et al.*, Infect. Immun. 63:2682, 1995); a 20 kDa membrane-associated lipoprotein (Kostrzynska *et al.*, J. Bact. 176:5938, 1994); a 30 kDa outer membrane protein (Bölin *et al.*, J. Clin. Microbiol. 33:381, 1995); a lactoferrin receptor (FR 2,724,936); and several porins, designated HopA, HopB, HopC, HopD, and HopE, which have molecular weights of 48-67 kDa (Exner *et al.*, Infect. Immun. 63:1567, 1995; Doig *et al.*, J. Bact. 177:5447, 1995). Some of these proteins have been proposed as potential vaccine antigens. In particular, urease is believed to be a vaccine candidate (WO 94/9823; WO 95/22987; WO 95/3824; Michetti *et al.*, Gastroenterology 107:1002, 1994). Nevertheless, it is thought that several antigens may ultimately be necessary in a vaccine.

Summary of the Invention

The invention provides polynucleotide molecules that encode *Helicobacter* polypeptides, designated GHPO 13, GHPO 73, GHPO 90, GHPO 107, GHPO 136, GHPO 191, GHPO 213, GHPO 240, GHPO 408, GHPO 411, GHPO 419, GHPO 431, GHPO 474, GHPO 591, GHPO 596, GHPO 699, GHPO 724, GHPO 730,

GHPO 761, GHPO 804, GHPO 805, GHPO 812, GHPO 879, GHPO 888, GHPO 986, GHPO 1056, GHPO 1081, GHPO 1100, GHPO 1140, GHPO 1148, GHPO 1200, GHPO 1212, GHPO 1258, GHPO 1263, GHPO 1273, GHPO 1284, GHPO 1299, GHPO 1327, GHPO 1346, GHPO 1378, GHPO 1412, GHPO 1443, GHPO 1466, GHPO 1476, GHPO 1536, GHPO 1559, GHPO 427, GHPO 1045, and GHPO 1262, which can be used, *e.g.*, in methods to prevent, treat, or diagnose *Helicobacter* infection. The polypeptides of the invention include those having the amino acid sequences shown in SEQ ID NOs:2-98 (even numbers), as well as mature forms of proteins having sequences shown in SEQ ID NOs:2-98 in their unprocessed forms, and fragments thereof.

Those skilled in the art will understand that the invention also includes polynucleotide molecules that encode mutants and derivatives of these polypeptides, which can result from the addition, deletion, or substitution of non-essential amino acids, as is described further below.

In addition to the polynucleotide molecules described above, the invention includes the corresponding polypeptides (*i.e.*, polypeptides encoded by the polynucleotide molecules of the invention, or fragments thereof), and monospecific antibodies that specifically bind to these polypeptides.

The present invention has many applications and includes expression cassettes, vectors, and cells transformed or transfected with the polynucleotides of the invention. Accordingly, the present invention provides (i) methods for producing polypeptides of the invention in recombinant host systems and related expression cassettes, vectors, and transformed or transfected cells; (ii) live vaccine vectors, such as pox virus, *Salmonella typhimurium*, and *Vibrio cholerae* vectors, that contain polynucleotides of the invention (such vaccine vectors being useful in, *e.g.*, methods for preventing or treating *Helicobacter* infection) in combination

with a diluent or carrier, and related pharmaceutical compositions and associated therapeutic and/or prophylactic methods; (iii) therapeutic and/or prophylactic methods involving administration of polynucleotide molecules, either in a naked form or formulated with a delivery vehicle, polypeptides or mixtures of polypeptides, or monospecific antibodies of the invention, and related pharmaceutical compositions; (iv) methods for detecting the presence of *Helicobacter* in biological samples, which can involve the use of polynucleotide molecules, monospecific antibodies, or polypeptides of the invention; and (v) methods for purifying polypeptides of the invention by antibody-based affinity chromatography.

Brief Description of the Drawings

Fig. 1A is a diagrammatic representation of transposon TnMax9, which is a derivative of the TnMax transposon system (Haas *et al.*, Gene 130:23-21, 1993). The mini-transposon carries the *blaM* gene, which is the β -lactamase gene lacking a promoter and a signal sequence, next to the inverted repeats (IR) and the M13 forward (M13-FP) and reverse (M13-RP1) primer binding sites. The resolution site (*res*) and an origin of replication (*ori_{rd}*) are located between the *BlaM* gene and the constitutive *cat_{cc}*-resistance gene. The transposase *tnpA* and resolvase *tnpR* genes are located outside of the mini-transposon and are under the control of the inducible P_{trc} promoter. The *lacIq* gene encodes the Lac repressor.

Fig. 1B is a diagrammatic representation of plasmid pMin2. pMin2 contains a multiple cloning site, the tetracycline resistance gene (*tet*), an origin of transfer (*oriT*), an origin of replication (*ori_{ColEI}*), a transcriptional terminator (*t_{rd}*),

and a weak, constitutive promoter (P_{iga}). *H. pylori* chromosome fragments were introduced into the *Bgl*III and *Cla*I sites of pMin2.

Detailed Description

Open reading frames (ORFs) encoding new, full length polypeptides, designated GHPO 13, GHPO 73, GHPO 90, GHPO 107, GHPO 136, GHPO 191, GHPO 213, GHPO 240, GHPO 408, GHPO 411, GHPO 419, GHPO 431, GHPO 474, GHPO 591, GHPO 596, GHPO 699, GHPO 724, GHPO 730, GHPO 761, GHPO 804, GHPO 805, GHPO 812, GHPO 879, GHPO 888, GHPO 986, GHPO 1056, GHPO 1081, GHPO 1100, GHPO 1140, GHPO 1148, GHPO 1200, GHPO 1212, GHPO 1258, GHPO 1263, GHPO 1273, GHPO 1284, GHPO 1299, GHPO 1327, GHPO 1346, GHPO 1378, GHPO 1412, GHPO 1443, GHPO 1466, GHPO 1476, GHPO 1536, GHPO 1559, GHPO 427, GHPO 1045, and GHPO 1262 have been identified in the *H. pylori* genome. These polypeptides can be used, for example, in vaccination methods for preventing or treating *Helicobacter* infection. Some of the new polypeptides are secreted polypeptides that can be produced in their mature forms (*i.e.*, as polypeptides that have been exported through class II or class III secretion pathways) or as precursors that include signal peptides, which can be removed in the course of excretion/secretion by cleavage at the N-terminal end of the mature form. (The cleavage site is located at the C-terminal end of the signal peptide, adjacent to the mature form.)

According to a first aspect of the invention, there are provided isolated polynucleotides that encode the precursor and mature forms of *Helicobacter* GHPO 13, GHPO 73, GHPO 90, GHPO 107, GHPO 136, GHPO 191, GHPO 213, GHPO 240, GHPO 408, GHPO 411, GHPO 419, GHPO 431, GHPO 474, GHPO 591, GHPO 596,

GHPO 699, GHPO 724, GHPO 730, GHPO 761, GHPO 804, GHPO 805, GHPO 812, GHPO 879, GHPO 888, GHPO 986, GHPO 1056, GHPO 1081, GHPO 1100, GHPO 1140, GHPO 1148, GHPO 1200, GHPO 1212, GHPO 1258, GHPO 1263, GHPO 1273, GHPO 1284, GHPO 1299, GHPO 1327, GHPO 1346, GHPO 1378, GHPO 1412, GHPO 1443, GHPO 1466, GHPO 1476, GHPO 1536, GHPO 1559, GHPO 427, GHPO 1045, and GHPO 1262. Polynucleotides designated GHPO 1424 (SEQ ID NO:99, ATG start codon at position 82) and GHPO 1736 (SEQ ID NO:100, ATG start codon at position 336) are also included in the invention.

An isolated polynucleotide of the invention encodes:

(i) a polypeptide having an amino acid sequence that is homologous to a *Helicobacter* amino acid sequence of a polypeptide, the *Helicobacter* amino acid sequence being selected from the group consisting of the amino acid sequences shown in SEQ ID NO:2 (GHPO 13), SEQ ID NO:4 (GHPO 73), SEQ ID NO:6 (GHPO 90), SEQ ID NO:8 (GHPO 107), SEQ ID NO:10 (GHPO 136), SEQ ID NO:12 (GHPO 191), SEQ ID NO:14 (GHPO 213), SEQ ID NO:16 (GHPO 240), SEQ ID NO:18 (GHPO 408), SEQ ID NO:20 (GHPO 411), SEQ ID NO:22 (GHPO 419), SEQ ID NO:24 (GHPO 431), SEQ ID NO:26 (GHPO 474), SEQ ID NO:28 (GHPO 591), SEQ ID NO:30 (GHPO 596), SEQ ID NO:32 (GHPO 699), SEQ ID NO:34 (GHPO 724), SEQ ID NO:36 (GHPO 730), SEQ ID NO:38 (GHPO 761), SEQ ID NO:40 (GHPO 804), SEQ ID NO:42 (GHPO 805), SEQ ID NO:44 (GHPO 812), SEQ ID NO:46 (GHPO 879), SEQ ID NO:48 (GHPO 888), SEQ ID NO:50 (GHPO 986), SEQ ID NO:52 (GHPO 1056), SEQ ID NO:54 (GHPO 1081), SEQ ID NO:56 (GHPO 1100), SEQ ID NO:58 (GHPO 1140), SEQ ID NO:60 (GHPO 1148), SEQ ID NO:62 (GHPO 1200), SEQ ID NO:64 (GHPO 1212), SEQ ID NO:66 (GHPO 1258), SEQ ID NO:68 (GHPO 1263), SEQ ID NO:70 (GHPO 1273), SEQ ID NO:72 (GHPO 1284), SEQ ID NO:74 (GHPO 1299), SEQ ID NO:76 (GHPO 1327), SEQ ID NO:78 (GHPO 1346), SEQ ID NO:80 (GHPO 1378), SEQ ID

NO:82 (GHPO 1412), SEQ ID NO:84 (GHPO 1443), SEQ ID NO:86 (GHPO 1466),
SEQ ID NO:88 (GHPO 1476), SEQ ID NO:90 (GHPO 1536), SEQ ID NO:92 (GHPO
1559), SEQ ID NO:94 (GHPO 427), SEQ ID NO:96 (GHPO 1045), and SEQ ID NO:98
(GHPO 1262); or

5 (ii) a derivative of the polypeptide.

In addition to the full-length polypeptides encoded by the
polynucleotides of the invention, as set forth above, polynucleotides included in
the invention can also encode polypeptides that lack signal sequences, as well as
other polypeptide or peptide fragments of the full-length polypeptides.

10 The term "isolated polynucleotide" is defined as a polynucleotide that is
removed from the environment in which it naturally occurs. For example, a
naturally-occurring DNA molecule present in the genome of a living bacteria or as
part of a gene bank is not isolated, but the same molecule, separated from the
remaining part of the bacterial genome, as a result of, *e.g.*, a cloning event
15 (amplification), is "isolated." Typically, an isolated DNA molecule is free from
DNA regions (*e.g.*, coding regions) with which it is immediately contiguous, at the
5' or 3' ends, in the naturally occurring genome. Such isolated polynucleotides can
be part of a vector or a composition and still be isolated, as such a vector or
composition is not part of its natural environment.

20 A polynucleotide of the invention can consist of RNA or DNA (*e.g.*,
cDNA, genomic DNA, or synthetic DNA), or modifications or combinations of
RNA or DNA. The polynucleotide can be double-stranded or single-stranded and,
if single-stranded, can be the coding (sense) strand or the non-coding (anti-sense)
strand. The sequences that encode polypeptides of the invention, as shown in any
25 of SEQ ID NOs:2-98 (even numbers), can be (a) the coding sequence as shown in

any of SEQ ID NOs:1-97 (odd numbers), 99, and 100; (b) a ribonucleotide sequence derived by transcription of (a); or (c) a different coding sequence that, as a result of the redundancy or degeneracy of the genetic code, encodes the same polypeptides as the polynucleotide molecules having the sequences illustrated in
5 any of SEQ ID NOs:1-97 (odd numbers), 99, and 100. The polypeptide can be one that is naturally secreted or excreted by, *e.g.*, *H. felis*, *H. mustelae*, *H. heilmanii*, or *H. pylori*.

By "polypeptide" or "protein" is meant any chain of amino acids, regardless of length or post-translational modification (*e.g.*, glycosylation or
10 phosphorylation). Both terms are used interchangeably in the present application.

By "homologous amino acid sequence" is meant an amino acid sequence that differs from an amino acid sequence shown in any of SEQ ID NOs:2-98 (even numbers), or an amino acid sequence encoded by the nucleotide sequence of any of SEQ ID NOs:1-97 (odd numbers), 99, and 100, by one or more non-
15 conservative amino acid substitutions, deletions, or additions located at positions at which they do not destroy the specific antigenicity of the polypeptide. Preferably, such a sequence is at least 75%, more preferably at least 80%, and most preferably at least 90% identical to an amino acid sequence shown in any of SEQ ID NOs:2-98 (even numbers).

20 Homologous amino acid sequences include sequences that are identical or substantially identical to an amino acid sequence as shown in any of SEQ ID NOs:2-98 (even numbers). By "amino acid sequence that is substantially identical" is meant a sequence that is at least 90%, preferably at least 95%, more preferably at least 97%, and most preferably at least 99% identical to an amino

acid sequence of reference and that differs from the sequence of reference, if at all, by a majority of conservative amino acid substitutions.

Conservative amino acid substitutions typically include substitutions among amino acids of the same class. These classes include, for example, amino acids having uncharged polar side chains, such as asparagine, glutamine, serine, threonine, and tyrosine; amino acids having basic side chains, such as lysine, arginine, and histidine; amino acids having acidic side chains, such as aspartic acid and glutamic acid; and amino acids having nonpolar side chains, such as glycine, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan, and cysteine.

Homology can be measured using sequence analysis software (*e.g.*, Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, WI 53705). Similar amino acid sequences are aligned to obtain the maximum degree of homology (*i.e.*, identity). To this end, it may be necessary to artificially introduce gaps into the sequence. Once the optimal alignment has been set up, the degree of homology (*i.e.*, identity) is established by recording all of the positions in which the amino acids of both sequences are identical, relative to the total number of positions.

Homologous polynucleotide sequences are defined in a similar way. Preferably, a homologous sequence is one that is at least 45%, more preferably at least 60%, and most preferably at least 85% identical to a coding sequence of any of SEQ ID NOs:1-97 (odd numbers), 99, and 100.

Polypeptides having a sequence homologous to any one of the sequences shown in SEQ ID NOs:2-98 (even numbers), include naturally-

occurring allelic variants, as well as mutants or any other non-naturally occurring variants that are analogous in terms of antigenicity, to a polypeptide having a sequence as shown in any one of SEQ ID NOs:2-98 (even numbers).

As is known in the art, an allelic variant is an alternate form of a polypeptide that is characterized as having a substitution, deletion, or addition of one or more amino acids that does not alter the biological function of the polypeptide. By "biological function" is meant a function of the polypeptide in the cells in which it naturally occurs, even if the function is not necessary for the growth or survival of the cells. For example, the biological function of a porin is to allow the entry into cells of compounds present in the extracellular medium. The biological function is distinct from the antigenic function. A polypeptide can have more than one biological function.

Allelic variants are very common in nature. For example, a bacterial species, *e.g.*, *H. pylori*, is usually represented by a variety of strains that differ from each other by minor allelic variations. Indeed, a polypeptide that fulfills the same biological function in different strains can have an amino acid sequence that is not identical in each of the strains. Such an allelic variation can be equally reflected at the polynucleotide level.

Support for the use of allelic variants of polypeptide antigens comes from, *e.g.*, studies of the *Helicobacter* urease antigen. The amino acid sequence of *Helicobacter* urease varies widely from species to species, yet cross-species protection occurs, indicating that the urease molecule, when used as an immunogen, is highly tolerant of amino acid variations. Even among different strains of the single species *H. pylori*, there are amino acid sequence variations.

For example, although the amino acid sequences of the UreA and UreB subunits of *H. pylori* and *H. felis* ureases differ from one another by 26.5% and 11.8%, respectively (Ferrero *et al.*, Molecular Microbiology 9(2):323-333, 1993), it has been shown that *H. pylori* urease protects mice from *H. felis* infection

5 (Michetti *et al.*, Gastroenterology 107:1002, 1994). In addition, it has been shown that the individual structural subunits of urease, UreA and UreB, which contain distinct amino acid sequences, are both protective antigens against *Helicobacter* infection (Michetti *et al.*, *supra*). Similarly, Cuenca *et al.* (Gastroenterology 110:1770, 1996) showed that therapeutic immunization of *H. mustelae*-infected
10 ferrets with *H. pylori* urease was effective at eradicating *H. mustelae* infection.

Further, several urease variants have been reported to be effective vaccine antigens, including, *e.g.*, recombinant UreA + UreB apoenzyme expressed from pORV142 (UreA and UreB sequences derived from *H. pylori* strain CPM630; Lee
15 *et al.*, J. Infect. Dis. 172:161, 1995); recombinant UreA + UreB apoenzyme expressed from pORV214 (UreA and UreB sequences differ from *H. pylori* strain CPM630 by one and two amino acid changes, respectively; Lee *et al.*, *supra*, 1995); a UreA-glutathione-S-transferase fusion protein (UreA sequence from *H. pylori* strain ATCC 43504; Thomas *et al.*, Acta Gastro-Enterologica Belgica 56:54, 1993); UreA + UreB holoenzyme purified from *H. pylori* strain
20 NCTC11637 (Marchetti *et al.*, Science 267:1655, 1995); a UreA-MBP fusion protein (UreA from *H. pylori* strain 85P; Ferrero *et al.*, Infection and Immunity 62:4981, 1994); a UreB-MBP fusion protein (UreB from *H. pylori* strain 85P; Ferrero *et al.*, *supra*); a UreA-MBP fusion protein (UreA from *H. felis* strain ATCC 49179; Ferrero *et al.*, *supra*); a UreB-MBP fusion protein (UreB from *H.*
25 *felis* strain ATCC 49179; Ferrero *et al.*, *supra*); and a 37 kDa fragment of UreB

containing amino acids 220-569 (Dore-Davin *et al.*, "A 37 kD fragment of UreB is sufficient to confer protection against *Helicobacter felis* infection in mice").

Finally, Thomas *et al.* (*supra*) showed that oral immunization of mice with crude sonicates of *H. pylori* protected mice from subsequent challenge with *H. felis*.

5 Polynucleotides, *e.g.*, DNA molecules, encoding allelic variants can easily be obtained by polymerase chain reaction (PCR) amplification of genomic bacterial DNA extracted by conventional methods. This involves the use of synthetic oligonucleotide primers matching sequences that are upstream and downstream of the 5' and 3' ends of the coding region. Suitable primers can be
10 designed based on the nucleotide sequence information provided in any of SEQ ID NOs:1-97 (odd numbers), 99, and 100. Typically, a primer consists of 10 to 40, preferably 15 to 25 nucleotides. It can also be advantageous to select primers containing C and G nucleotides in proportions sufficient to ensure efficient hybridization, *e.g.*, an amount of C and G nucleotides of at least 40%, preferably
15 50%, of the total nucleotide amount. Those skilled in the art can readily design primers that can be used to isolate the polynucleotides of the invention from different *Helicobacter* strains. Experimental conditions for carrying out PCR can readily be determined by one skilled in the art and an illustration of carrying out PCR is provided in Example 2. As is well known in the art, restriction
20 endonuclease recognition sites that contain, typically, 4 to 6 nucleotides (for example, the sequences 5'-GGATCC-3' (*Bam*HI) or 5'-CTCGAG-3' (*Xho*I)), can be included on the 5' ends of the primers. Restriction sites can be selected by those skilled in the art so that the amplified DNA can be conveniently cloned into an appropriately digested vector, such as a plasmid.

Useful homologs that do not occur naturally can be designed using known methods for identifying regions of an antigen that are likely to be tolerant of amino acid sequence changes and/or deletions. For example, sequences of the antigen from different species can be compared to identify conserved sequences.

5 Polypeptide derivatives that are encoded by polynucleotides of the invention include, *e.g.*, fragments, polypeptides having large internal deletions derived from full-length polypeptides, and fusion proteins. Polypeptide fragments of the invention can be derived from a polypeptide having a sequence homologous to any of the sequences of SEQ ID NOs:2-98 (even numbers), to the extent that the
10 fragments retain the substantial antigenicity of the parent polypeptide (specific antigenicity). Polypeptide derivatives can also be constructed by large internal deletions that remove a substantial part of the parent polypeptide, while retaining specific antigenicity. Generally, polypeptide derivatives should be about at least 12 amino acids in length to maintain antigenicity. Advantageously, they can be at
15 least 20 amino acids, preferably at least 50 amino acids, more preferably at least 75 amino acids, and most preferably at least 100 amino acids in length.

Useful polypeptide derivatives, *e.g.*, polypeptide fragments, can be designed using computer-assisted analysis of amino acid sequences in order to identify sites in protein antigens having potential as surface-exposed, antigenic
20 regions (Hughes *et al.*, Infect. Immun. 60(9):3497, 1992). For example, the Laser Gene Program from DNA Star can be used to obtain hydrophilicity, antigenic index, and intensity index plots for the polypeptides of the invention. This program can also be used to obtain information about homologies of the polypeptides with known protein motifs. One skilled in the art can readily use the
25 information provided in such plots to select peptide fragments for use as vaccine

antigens. For example, fragments spanning regions of the plots in which the antigenic index is relatively high can be selected. One can also select fragments spanning regions in which both the antigenic index and the intensity plots are relatively high. Fragments containing conserved sequences, particularly hydrophilic conserved sequences, can also be selected.

Polypeptide fragments and polypeptides having large internal deletions can be used for revealing epitopes that are otherwise masked in the parent polypeptide and that may be of importance for inducing a protective T cell-dependent immune response. Deletions can also remove immunodominant regions of high variability among strains.

It is an accepted practice in the field of immunology to use fragments and variants of protein immunogens as vaccines, as all that is required to induce an immune response to a protein is a small (*e.g.*, 8 to 10 amino acids) immunogenic region of the protein. This has been done for a number of vaccines against pathogens other than *Helicobacter*. For example, short synthetic peptides corresponding to surface-exposed antigens of pathogens such as murine mammary tumor virus (peptide containing 11 amino acids; Dion *et al.*, Virology 179:474-477, 1990), Semliki Forest virus (peptide containing 16 amino acids; Snijders *et al.*, J. Gen. Virol. 72:557-565, 1991), and canine parvovirus (2 overlapping peptides, each containing 15 amino acids; Langeveld *et al.*, Vaccine 12(15):1473-1480, 1994) have been shown to be effective vaccine antigens against their respective pathogens.

Polynucleotides encoding polypeptide fragments and polypeptides having large internal deletions can be constructed using standard methods (see, *e.g.*, Ausubel *et al.*, *Current Protocols in Molecular Biology*, John Wiley & Sons

Inc., 1994), for example, by PCR, including inverse PCR, by restriction enzyme treatment of the cloned DNA molecules, or by the method of Kunkel *et al.* (Proc. Natl. Acad. Sci. USA 82:448, 1985; biological material available at Stratagene).

A polypeptide derivative can also be produced as a fusion polypeptide that contains a polypeptide or a polypeptide derivative of the invention fused, *e.g.*, at the N- or C-terminal end, to any other polypeptide (hereinafter referred to as a peptide tail). Such a product can be easily obtained by translation of a genetic fusion, *i.e.*, a hybrid gene. Vectors for expressing fusion polypeptides are commercially available, and include the pMal-c2 or pMal-p2 systems of New England Biolabs, in which the peptide tail is a maltose binding protein, the glutathione-S-transferase system of Pharmacia, or the His-Tag system available from Novagen. These and other expression systems provide convenient means for further purification of polypeptides and derivatives of the invention.

Another particular example of fusion polypeptides included in invention includes a polypeptide or polypeptide derivative of the invention fused to a polypeptide having adjuvant activity, such as, *e.g.*, subunit B of either cholera toxin or *E. coli* heat-labile toxin. Several possibilities can be used for producing such fusion proteins. First, the polypeptide of the invention can be fused to the N-terminal end or, preferably, to the C-terminal end of the polypeptide having adjuvant activity. Second, a polypeptide fragment of the invention can be fused within the amino acid sequence of the polypeptide having adjuvant activity. Spacer sequences can also be included, if desired.

As stated above, the polynucleotides of the invention encode *Helicobacter* polypeptides in precursor or mature form. They can also encode hybrid precursors containing heterologous signal peptides, which can mature into

polypeptides of the invention. By "heterologous signal peptide" is meant a signal peptide that is not found in the naturally-occurring precursor of a polypeptide of the invention.

A polynucleotide of the invention hybridizes, preferably under stringent
5 conditions, to a polynucleotide having a sequence as shown in any of SEQ ID
NOs:1-97 (odd numbers), 99, and 100. Hybridization procedures are, *e.g.*,
described by Ausubel *et al.* (*supra*); Silhavy *et al.* (*Experiments with Gene
Fusions*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York,
1984); and Davis *et al.* (*A Manual for Genetic Engineering: Advanced Bacterial
10 Genetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York,
1980). Important parameters that can be considered for optimizing hybridization
conditions are reflected in the following formula, which facilitates calculation of
the melting temperature (T_m), which is the temperature above which two
complementary DNA strands separate from one another (Casey *et al.*, Nucl. Acid
15 Res. 4:1539, 1977): $T_m = 81.5 + 0.5 \times (\% \text{ G+C}) + 1.6 \log (\text{positive ion}$
 $\text{concentration}) - 0.6 \times (\% \text{ formamide})$. Under appropriate stringency conditions,
hybridization temperature (T_h) is approximately 20 to 40°C, 20 to 25°C, or,
preferably, 30 to 40°C below the calculated T_m . Those skilled in the art will
understand that optimal temperature and salt conditions can be readily determined
20 empirically in preliminary experiments using conventional procedures. For
example, stringent conditions can be achieved, both for pre-hybridizing and
hybridizing incubations, (i) within 4-16 hours at 42°C, in 6 x SSC containing
50% formamide or (ii) within 4-16 hours at 65°C in an aqueous 6 x SSC solution
(1 M NaCl, 0.1 M sodium citrate (pH 7.0)). For polynucleotides containing 30 to
25 600 nucleotides, the above formula is used and then is corrected by subtracting

(600/polynucleotide size in base pairs). Stringency conditions are defined by a T_h that is 5 to 10°C below T_m .

Hybridization conditions with oligonucleotides shorter than 20-30 bases do not precisely follow the rules set forth above. In such cases, the formula for calculating the T_m is as follows: $T_m = 4 \times (G+C) + 2 (A+T)$. For example, an 18 nucleotide fragment of 50% G+C would have an approximate T_m of 54°C.

A polynucleotide molecule of the invention, containing RNA, DNA, or modifications or combinations thereof, can have various applications. For example, a polynucleotide molecule can be used (i) in a process for producing the encoded polypeptide in a recombinant host system, (ii) in the construction of vaccine vectors such as poxviruses, which are further used in methods and compositions for preventing and/or treating *Helicobacter* infection, (iii) as a vaccine agent, in a naked form or formulated with a delivery vehicle and, (iv) in the construction of attenuated *Helicobacter* strains that can over-express a polynucleotide of the invention or express it in a non-toxic, mutated form.

According to a second aspect of the invention, there is therefore provided (i) an expression cassette containing a polynucleotide molecule of the invention placed under the control of elements (*e.g.*, a promoter) required for expression; (ii) an expression vector containing an expression cassette of the invention; (iii) a procaryotic or eucaryotic cell transformed or transfected with an expression cassette and/or vector of the invention, as well as (iv) a process for producing a polypeptide or polypeptide derivative encoded by a polynucleotide of the invention, which involves culturing a procaryotic or eucaryotic cell transformed or transfected with an expression cassette and/or vector of the invention, under conditions that allow expression of the polynucleotide molecule

of the invention and, recovering the encoded polypeptide or polypeptide derivative from the cell culture.

A recombinant expression system can be selected from procaryotic and eucaryotic hosts. Eucaryotic hosts include, for example, yeast cells (*e.g.*,
5 *Saccharomyces cerevisiae* or *Pichia Pastoris*), mammalian cells (*e.g.*, COS1, NIH3T3, or JEG3 cells), arthropods cells (*e.g.*, *Spodoptera frugiperda* (SF9) cells), and plant cells. Preferably, a procaryotic host such as *E. coli* is used. Bacterial and eucaryotic cells are available from a number of different sources that are known to those skilled in the art, *e.g.*, the American Type Culture Collection
10 (ATCC; Rockville, Maryland).

The choice of the expression cassette will depend on the host system selected, as well as the features desired for the expressed polypeptide. For example, it may be useful to produce a polypeptide of the invention in a particular lipidated form or any other form. Typically, an expression cassette includes a
15 constitutive or inducible promoter that is functional in the selected host system; a ribosome binding site; a start codon (ATG); if necessary, a region encoding a signal peptide, *e.g.*, a lipidation signal peptide; a polynucleotide molecule of the invention; a stop codon; and, optionally, a 3' terminal region (translation and/or transcription terminator). The signal peptide-encoding region is adjacent to the
20 polynucleotide of the invention and is placed in the proper reading frame. The signal peptide-encoding region can be homologous or heterologous to the polynucleotide molecule encoding the mature polypeptide and it can be specific to the secretion apparatus of the host used for expression. The open reading frame constituted by the polynucleotide molecule of the invention, alone or together with
25 the signal peptide, is placed under the control of the promoter so that transcription

and translation occur in the host system. Promoters and signal peptide-encoding regions are widely known and available to those skilled in the art and include, for example, the promoter of *Salmonella typhimurium* (and derivatives) that is inducible by arabinose (promoter araB) and is functional in Gram-negative bacteria such as *E. coli* (U.S. Patent No. 5,028,530; Cagnon *et al.*, Protein Engineering 4(7):843, 1991); the promoter of the bacteriophage T7 RNA polymerase gene, which is functional in a number of *E. coli* strains expressing T7 polymerase (U.S. Patent No. 4,952,496); the OspA lipidation signal peptide; and RlpB lipidation signal peptide (Takase *et al.*, J. Bact. 169:5692, 1987).

The expression cassette is typically part of an expression vector, which is selected for its ability to replicate in the chosen expression system. Expression vectors (*e.g.*, plasmids or viral vectors) can be chosen from, for example, those described in Pouwels *et al.* (*Cloning Vectors: A Laboratory Manual*, 1985, Supp. 1987) and can be purchased from various commercial sources. Methods for transforming or transfecting host cells with expression vectors are well known in the art and will depend on the host system selected, as described in Ausubel *et al.* (*supra*).

Upon expression, a recombinant polypeptide of the invention (or a polypeptide derivative) is produced and remains in the intracellular compartment, is secreted/excreted in the extracellular medium or in the periplasmic space, or is embedded in the cellular membrane. The polypeptide can then be recovered in a substantially purified form from the cell extract or from the supernatant after centrifugation of the cell culture. Typically, the recombinant polypeptide can be purified by antibody-based affinity purification or by any other method known to a person skilled in the art, such as by genetic fusion to a small affinity-binding

domain. Antibody-based affinity purification methods are also available for purifying a polypeptide of the invention extracted from a *Helicobacter* strain. Antibodies useful for immunoaffinity purification of the polypeptides of the invention can be obtained using methods described below.

5 Polynucleotides of the invention can also be used in DNA vaccination methods, using either a viral or bacterial host as gene delivery vehicle (live vaccine vector) or administering the gene in a free form, *e.g.*, inserted into a plasmid. Therapeutic or prophylactic efficacy of a polynucleotide of the invention can be evaluated as is described below.

10 Accordingly, in a third aspect of the invention, there is provided (i) a vaccine vector such as a poxvirus, containing a polynucleotide molecule of the invention placed under the control of elements required for expression; (ii) a composition of matter containing a vaccine vector of the invention, together with a diluent or carrier; (iii) a pharmaceutical composition containing a therapeutically
15 or prophylactically effective amount of a vaccine vector of the invention; (iv) a method for inducing an immune response against *Helicobacter* in a mammal (*e.g.*, a human; alternatively, the method can be used in veterinary applications for treating or preventing *Helicobacter* infection of animals, *e.g.*, cats or birds), which involves administering to the mammal an immunogenically effective amount of a
20 vaccine vector of the invention to elicit an immune response, *e.g.*, a protective or therapeutic immune response to *Helicobacter*; and (v) a method for preventing and/or treating a *Helicobacter* (*e.g.*, *H. pylori*, *H. felis*, *H. mustelae*, or *H. heilmanii*) infection, which involves administering a prophylactic or therapeutic amount of a vaccine vector of the invention to an individual in need. Additionally,
25 the third aspect of the invention encompasses the use of a vaccine vector of the

invention in the preparation of a medicament for preventing and/or treating *Helicobacter* infection.

A vaccine vector of the invention can express one or several polypeptides or derivatives of the invention, as well as at least one additional *Helicobacter* antigen such as a urease apoenzyme or a subunit, fragment, homolog, mutant, or derivative thereof. In addition, it can express a cytokine, such as interleukin-2 (IL-2) or interleukin-12 (IL-12), that enhances the immune response. Thus, a vaccine vector can include an additional polynucleotide molecules encoding, *e.g.*, urease subunit A, B, or both, or a cytokine, placed under the control of elements required for expression in a mammalian cell.

Alternatively, a composition of the invention can include several vaccine vectors, each of which being capable of expressing a polypeptide or derivative of the invention. A composition can also contain a vaccine vector capable of expressing an additional *Helicobacter* antigen such as urease apoenzyme, a subunit, fragment, homolog, mutant, or derivative thereof, or a cytokine such as IL-2 or IL-12.

In vaccination methods for treating or preventing infection in a mammal, a vaccine vector of the invention can be administered by any conventional route in use in the vaccine field, for example, to a mucosal (*e.g.*, ocular, intranasal, oral, gastric, pulmonary, intestinal, rectal, vaginal, or urinary tract) surface or *via* a parenteral (*e.g.*, subcutaneous, intradermal, intramuscular, intravenous, or intraperitoneal) route. Preferred routes depend upon the choice of the vaccine vector. The administration can be achieved in a single dose or repeated at intervals. The appropriate dosage depends on various parameters that are understood by those skilled in the art, such as the nature of the vaccine vector

itself, the route of administration, and the condition of the mammal to be vaccinated (*e.g.*, the weight, age, and general health of the mammal).

Live vaccine vectors that can be used in the invention include viral vectors, such as adenoviruses and poxviruses, as well as bacterial vectors, *e.g.*,
5 *Shigella*, *Salmonella*, *Vibrio cholerae*, *Lactobacillus*, Bacille bilié de Calmette-Guérin (BCG), and *Streptococcus*. An example of an adenovirus vector, as well as a method for constructing an adenovirus vector capable of expressing a polynucleotide molecule of the invention, is described in U.S. Patent No. 4,920,209. Poxvirus vectors that can be used in the invention include, *e.g.*,
10 vaccinia and canary pox viruses, which are described in U.S. Patent No. 4,722,848 and U.S. Patent No. 5,364,773, respectively (also see, *e.g.*, Tartaglia *et al.*, Virology 188:217, 1992, for a description of a vaccinia virus vector, and Taylor *et al.*, Vaccine 13:539, 1995, for a description of a canary poxvirus vector). Poxvirus vectors capable of expressing a polynucleotide of the invention can be obtained by
15 homologous recombination, as described in Kieny *et al.* (Nature 312:163, 1984) so that the polynucleotide of the invention is inserted in the viral genome under appropriate conditions for expression in mammalian cells. Generally, the dose of viral vector vaccine, for therapeutic or prophylactic use, can be from about 1×10^4 to about 1×10^{11} , advantageously from about 1×10^7 to about 1×10^{10} , or, preferably,
20 from about 1×10^7 to about 1×10^9 plaque-forming units per kilogram. Preferably, viral vectors are administered parenterally, for example, in 3 doses that are 4 weeks apart. Those skilled in the art will recognize that it is preferable to avoid adding a chemical adjuvant to a composition containing a viral vector of the invention and thereby minimizing the immune response to the viral vector itself.

Non-toxicogenic *Vibrio cholerae* mutant strains that can be used in live oral vaccines are described by Mekalanos *et al.* (Nature 306:551, 1983) and in U.S. Patent No. 4,882,278 (strain in which a substantial amount of the coding sequence of each of the two *ctxA* alleles has been deleted so that no functional *cholerae* toxin is produced); WO 92/11354 (strain in which the *irgA* locus is inactivated by mutation; this mutation can be combined in a single strain with *ctxA* mutations); and WO 94/1533 (deletion mutant lacking functional *ctxA* and *attRS1* DNA sequences). These strains can be genetically engineered to express heterologous antigens, as described in WO 94/19482. An effective vaccine dose of a *V. cholerae* strain capable of expressing a polypeptide or polypeptide derivative encoded by a polynucleotide molecule of the invention can contain, *e.g.*, about 1×10^5 to about 1×10^9 , preferably about 1×10^6 to about 1×10^8 viable bacteria in an appropriate volume for the selected route of administration. Preferred routes of administration include all mucosal routes, but, most preferably, these vectors are administered intranasally or orally.

Attenuated *Salmonella typhimurium* strains, genetically engineered for recombinant expression of heterologous antigens, and their use as oral vaccines, are described by Nakayama *et al.* (Bio/Technology 6:693, 1988) and in WO 92/11361. Preferred routes of administration for these vectors include all mucosal routes. Most preferably, the vectors are administered intranasally or orally.

Others bacterial strains useful as vaccine vectors are described by High *et al.* (EMBO 11:1991, 1992) and Sizemore *et al.* (Science 270:299, 1995; *Shigella flexneri*); Medaglini *et al.* (Proc. Natl. Acad. Sci. USA 92:6868, 1995; (*Streptococcus gordonii*); Flynn (Cell. Mol. Biol. 40 (suppl. I):31, 1194), and in WO 88/6626, WO 90/0594, WO 91/13157, WO 92/1796, and WO 92/21376

(Bacille Calmette Guerin). In bacterial vectors, a polynucleotide of the invention can be inserted into the bacterial genome or it can remain in a free state, for example, carried on a plasmid.

An adjuvant can also be added to a composition containing a bacterial
5 vector vaccine. A number of adjuvants that can be used are known to those skilled in the art. For example, preferred adjuvants can be selected from the list provided below.

According to a fourth aspect of the invention, there is also provided (i) a composition of matter containing a polynucleotide of the invention, together with a
10 diluent or carrier; (ii) a pharmaceutical composition containing a therapeutically or prophylactically effective amount of a polynucleotide of the invention; (iii) a method for inducing an immune response against *Helicobacter*, in a mammal, by administering to the mammal an immunogenically effective amount of a polynucleotide of the invention to elicit an immune response, *e.g.*, a protective
15 immune response to *Helicobacter*; and (iv) a method for preventing and/or treating a *Helicobacter* (*e.g.*, *H. pylori*, *H. felis*, *H. mustelae*, or *H. heilmanii*) infection, by administering a prophylactic or therapeutic amount of a polynucleotide of the invention to an individual in need of such treatment. Additionally, the fourth aspect of the invention encompasses the use of a polynucleotide of the invention in
20 the preparation of a medicament for preventing and/or treating *Helicobacter* infection. The fourth aspect of the invention preferably includes the use of a polynucleotide molecule placed under conditions for expression in a mammalian cell, *e.g.*, in a plasmid that is unable to replicate in mammalian cells and to substantially integrate into a mammalian genome.

Polynucleotides (for example, DNA or RNA molecules) of the invention can also be administered as such to a mammal as a vaccine. When a DNA molecule of the invention is used, it can be in the form of a plasmid that is unable to replicate in a mammalian cell and unable to integrate into the mammalian genome. Typically, a DNA molecule is placed under the control of a promoter suitable for expression in a mammalian cell. The promoter can function ubiquitously or tissue-specifically. Examples of non-tissue specific promoters include the early Cytomegalovirus (CMV) promoter (U.S. Patent No. 4,168,062) and the Rous Sarcoma Virus promoter (Norton *et al.*, Molec. Cell Biol. 5:281, 1985). The desmin promoter (Li *et al.*, Gene 78:243, 1989; Li *et al.*, J. Biol. Chem. 266:6562, 1991; Li *et al.*, J. Biol. Chem. 268:10403, 1993) is tissue-specific and drives expression in muscle cells. More generally, useful promoters and vectors are described, *e.g.*, in WO 94/21797 and by Hartikka *et al.* (Human Gene Therapy 7:1205, 1996).

For DNA/RNA vaccination, the polynucleotide of the invention can encode a precursor or a mature form of a polypeptide of the invention. When it encodes a precursor form, the precursor sequence can be homologous or heterologous. In the latter case, a eucaryotic leader sequence can be used, such as the leader sequence of the tissue-type plasminogen factor (tPA).

A composition of the invention can contain one or several polynucleotides of the invention. It can also contain at least one additional polynucleotide encoding another *Helicobacter* antigen, such as urease subunit A, B, or both, or a fragment, derivative, mutant, or analog thereof. A polynucleotide encoding a cytokine, such as interleukin-2 (IL-2) or interleukin-12 (IL-12), can also be added to the composition so that the immune response is enhanced. These

additional polynucleotides are placed under appropriate control for expression. Advantageously, DNA molecules of the invention and/or additional DNA molecules to be included in the same composition are carried in the same plasmid.

Standard methods can be used in the preparation of therapeutic polynucleotides of the invention. For example, a polynucleotide can be used in a naked form, free of any delivery vehicles, such as anionic liposomes, cationic lipids, microparticles, *e.g.*, gold microparticles, precipitating agents, *e.g.*, calcium phosphate, or any other transfection-facilitating agent. In this case, the polynucleotide can be simply diluted in a physiologically acceptable solution, such as sterile saline or sterile buffered saline, with or without a carrier. When present, the carrier preferably is isotonic, hypotonic, or weakly hypertonic, and has a relatively low ionic strength, such as provided by a sucrose solution, *e.g.*, a solution containing 20% sucrose.

Alternatively, a polynucleotide can be associated with agents that assist in cellular uptake. It can be, *e.g.*, (i) complemented with a chemical agent that modifies cellular permeability, such as bupivacaine (see, *e.g.*, WO 94/16737), (ii) encapsulated into liposomes, or (iii) associated with cationic lipids or silica, gold, or tungsten microparticles.

Anionic and neutral liposomes are well-known in the art (see, *e.g.*, *Liposomes: A Practical Approach*, RPC New Ed, IRL Press, 1990, for a detailed description of methods for making liposomes) and are useful for delivering a large range of products, including polynucleotides.

Cationic lipids can also be used for gene delivery. Such lipids include, for example, Lipofectin™, which is also known as DOTMA (N-[1-(2,3-dioleoyloxy)propyl]-N,N,N-trimethylammonium chloride), DOTAP (1,2-

bis(oleyloxy)-3-(trimethylammonio)propane), DDAB (dimethyldioctadecylammonium bromide), DOGS (dioctadecylamidologlycyl spermine), and cholesterol derivatives. A description of these cationic lipids can be found in EP 187,702, WO 90/11092, U.S. Patent No. 5,283,185, WO 91/15501, 5 WO 95/26356, and U.S. Patent No. 5,527,928. Cationic lipids for gene delivery are preferably used in association with a neutral lipid such as DOPE (dioleoyl phosphatidylethanolamine; WO 90/11092). Other transfection-facilitating compounds can be added to a formulation containing cationic liposomes. A number of them are described in, *e.g.*, WO 93/18759, WO 93/19768, WO 10 94/25608, and WO 95/2397. They include, *e.g.*, spermine derivatives useful for facilitating the transport of DNA through the nuclear membrane (see, for example, WO 93/18759) and membrane-permeabilizing compounds such as GALA, Gramicidine S, and cationic bile salts (see, for example, WO 93/19768).

Gold or tungsten microparticles can also be used for gene delivery, as 15 described in WO 91/359, WO 93/17706, and by Tang *et al.* (Nature 356:152, 1992). In this case, the microparticle-coated polynucleotides can be injected *via* intradermal or intraepidermal routes using a needleless injection device ("gene gun"), such as those described in U.S. Patent No. 4,945,050, U.S. Patent No. 5,015,580, and WO 94/24263.

20 The amount of DNA to be used in a vaccine recipient depends, *e.g.*, on the strength of the promoter used in the DNA construct, the immunogenicity of the expressed gene product, the condition of the mammal intended for administration (*e.g.*, the weight, age, and general health of the mammal), the mode of administration, and the type of formulation. In general, a therapeutically or 25 prophylactically effective dose from about 1 μ g to about 1 mg, preferably, from

about 10 μg to about 800 μg , and, more preferably, from about 25 μg to about 250 μg , can be administered to human adults. The administration can be achieved in a single dose or repeated at intervals.

The route of administration can be any conventional route used in the vaccine field. As general guidance, a polynucleotide of the invention can be administered *via* a mucosal surface, *e.g.*, an ocular, intranasal, pulmonary, oral, intestinal, rectal, vaginal, or urinary tract surface, or *via* a parenteral route, *e.g.*, by an intravenous, subcutaneous, intraperitoneal, intradermal, intraepidermal, or intramuscular route. The choice of administration route will depend on, *e.g.*, the formulation that is selected. A polynucleotide formulated in association with bupivacaine is advantageously administered into muscle. When a neutral or anionic liposome or a cationic lipid, such as DOTMA, is used, the formulation can be advantageously injected *via* intravenous, intranasal (for example, by aerosolization), intramuscular, intradermal, and subcutaneous routes. A polynucleotide in a naked form can advantageously be administered *via* the intramuscular, intradermal, or subcutaneous routes. Although not absolutely required, such a composition can also contain an adjuvant. A systemic adjuvant that does not require concomitant administration in order to exhibit an adjuvant effect is preferable.

The sequence information provided in the present application enables the design of specific nucleotide probes and primers that can be used in diagnostic methods. Accordingly, in a fifth aspect of the invention, there is provided a nucleotide probe or primer having a sequence found in, or derived by degeneracy of the genetic code from, a sequence shown in any of SEQ ID NOs:1-97 (odd numbers), 99, and 100.

The term "probe" as used in the present application refers to DNA (preferably single stranded) or RNA molecules (or modifications or combinations thereof) that hybridize under the stringent conditions, as defined above, to polynucleotide molecules having sequences homologous to any of those shown in SEQ ID NOs:1-97 (odd numbers), 99, and 100, or to a complementary or anti-sense sequence of any of those shown in SEQ ID NOs:1-97 (odd numbers), 99, and 100. Generally, probes are significantly shorter than the full-length sequences shown in SEQ ID NOs:1-97 (odd numbers), 99, and 100. For example, they can contain from about 5 to about 100, preferably from about 10 to about 80 nucleotides. In particular, probes have sequences that are at least 75%, preferably at least 85%, more preferably 95% homologous to a portion of a sequence as shown in any of SEQ ID NOs:1-97 (odd numbers), 99, and 100 or a sequence complementary to any of such sequences.

Probes can contain modified bases, such as inosine, methyl-5-deoxycytidine, deoxyuridine, dimethylamino-5-deoxyuridine, or diamino-2, 6-purine. Sugar or phosphate residues can also be modified or substituted. For example, a deoxyribose residue can be replaced by a polyamide (Nielsen *et al.*, Science 254:1497, 1991) and phosphate residues can be replaced by ester groups such as diphosphate, alkyl, arylphosphonate, and phosphorothioate esters. In addition, the 2'-hydroxyl group on ribonucleotides can be modified by addition of, *e.g.*, alkyl groups.

Probes of the invention can be used in diagnostic tests, or as capture or detection probes. Such capture probes can be immobilized on solid supports, directly or indirectly, by covalent means or by passive adsorption. A detection probe can be labeled by a detectable label, for example a label selected from

radioactive isotopes; enzymes, such as peroxidase and alkaline phosphatase; enzymes that are able to hydrolyze a chromogenic, fluorogenic, or luminescent substrate; compounds that are chromogenic, fluorogenic, or luminescent; nucleotide base analogs; and biotin.

5 Probes of the invention can be used in any conventional hybridization method, such as in dot blot methods (Maniatis *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1982), Southern blot methods (Southern, J. Mol. Biol. 98:503, 1975), northern blot methods (identical to Southern blot to the exception that RNA is used
10 as a target), or a sandwich method (Dunn *et al.*, Cell 12:23, 1977). As is known in the art, the latter technique involves the use of a specific capture probe and a specific detection probe that have nucleotide sequences that are at least partially different from each other.

Primers used in the invention usually contain about 10 to 40 nucleotides
15 and are used to initiate enzymatic polymerization of DNA in an amplification process (*e.g.*, PCR), an elongation process, or a reverse transcription method. In a diagnostic method involving PCR, the primers can be labeled.

Thus, the invention also encompasses (i) a reagent containing a probe of the invention for detecting and/or identifying the presence of *Helicobacter* in a
20 biological material; (ii) a method for detecting and/or identifying the presence of *Helicobacter* in a biological material, in which (a) a sample is recovered or derived from the biological material, (b) DNA or RNA is extracted from the material and denatured, and (c) the sample is exposed to a probe of the invention, for example, a capture probe, a detection probe, or both, under stringent hybridization conditions,
25 so that hybridization is detected; and (iii) a method for detecting and/or identifying

the presence of *Helicobacter* in a biological material, in which (a) a sample is recovered or derived from the biological material, (b) DNA is extracted therefrom, (c) the extracted DNA is contacted with at least one, or, preferably two, primers of the invention, and amplified by the polymerase chain reaction, and (d) an
5 amplified DNA molecule is produced.

As mentioned above, polypeptides that can be produced by expression of the polynucleotides of the invention can be used as vaccine antigens. Accordingly, a sixth aspect of the invention features a substantially purified polypeptide or polypeptide derivative having an amino acid sequence encoded by a
10 polynucleotide of the invention.

A "substantially purified polypeptide" is defined as a polypeptide that is separated from the environment in which it naturally occurs and/or a polypeptide that is free of most of the other polypeptides that are present in the environment in which it was synthesized. The polypeptides of the invention can be purified from
15 a natural source, such as a *Helicobacter* strain, or can be produced using recombinant methods.

Homologous polypeptides or polypeptide derivatives encoded by polynucleotides of the invention can be screened for specific antigenicity by testing cross-reactivity with an antiserum raised against a polypeptide having an
20 amino acid sequence as shown in any of SEQ ID NOs:2-98 (even numbers). Briefly, a monospecific hyperimmune antiserum can be raised against a purified reference polypeptide as such or as a fusion polypeptide, for example, an expression product of MBP, GST, or His-tag systems, or a synthetic peptide predicted to be antigenic. The homologous polypeptide or derivative that is
25 screened for specific antigenicity can be produced as such or as a fusion

polypeptide. In the latter case, and if the antiserum is also raised against a fusion polypeptide, two different fusion systems are employed. Specific antigenicity can be determined using a number of methods, including Western blot (Towbin *et al.*, Proc. Natl. Acad. Sci. USA 76:4350, 1979), dot blot, and ELISA methods, as
5 described below.

In a Western blot assay, the product to be screened, either as a purified preparation or a total *E. coli* extract, is fractionated by SDS-PAGE, as described, for example, by Laemmli (Nature 227:680, 1970). After being transferred to a filter, such as a nitrocellulose membrane, the material is incubated with the
10 monospecific hyperimmune antiserum, which is diluted in a range of dilutions from about 1:50 to about 1:5000, preferably from about 1:100 to about 1:500. Specific antigenicity is shown once a band corresponding to the product exhibits reactivity at any of the dilutions in the range.

In an ELISA assay, the product to be screened can be used as the
15 coating antigen. A purified preparation is preferred, but a whole cell extract can also be used. Briefly, about 100 μ L of a preparation of about 10 μ g protein/mL is distributed into wells of a 96-well ELISA plate. The plate is incubated for about 2 hours at 37°C, then overnight at 4°C. The plate is washed with phosphate buffer saline (PBS) containing 0.05% Tween 20 (PBS/Tween buffer) and the wells are
20 saturated with 250 μ L PBS containing 1% bovine serum albumin (BSA), to prevent non-specific antibody binding. After 1 hour of incubation at 37°C, the plate is washed with PBS/Tween buffer. The antiserum is serially diluted in PBS/Tween buffer containing 0.5% BSA, and 100 μ L dilutions are added to each well. The plate is incubated for 90 minutes at 37°C, washed, and evaluated using
25 standard methods. For example, a goat anti-rabbit peroxidase conjugate can be

added to the wells when the specific antibodies used were raised in rabbits.

Incubation is carried out for about 90 minutes at 37°C and the plate is washed.

The reaction is developed with the appropriate substrate and the reaction is

measured by colorimetry (absorbance measured spectrophotometrically). Under

5 these experimental conditions, a positive reaction is shown once an O.D. value of 1.0 is detected with a dilution of at least about 1:50, preferably of at least about 1:500.

In a dot blot assay, a purified product is preferred, although a whole cell extract can be used. Briefly, a solution of the product at a concentration of about

10 100 µg/mL is serially diluted two-fold with 50 mM Tris-HCl (pH 7.5). One hundred µL of each dilution is applied to a filter, such as a 0.45 µm nitrocellulose membrane, set in a 96-well dot blot apparatus (Biorad). The buffer is removed by applying vacuum to the system. Wells are washed by addition of 50 mM Tris-HCl (pH 7.5) and the membrane is air-dried. The membrane is saturated in blocking
15 buffer (50 mM Tris-HCl (pH 7.5), 0.15 M NaCl, 10 g/L skim milk) and incubated with an antiserum diluted from about 1:50 to about 1:5000, preferably about 1:500. The reaction is detected using standard methods. For example, a goat anti-rabbit peroxidase conjugate can be added to the wells when rabbit antibodies are used.

Incubation is carried out for about 90 minutes at 37°C and the blot is washed. The

20 reaction is developed with the appropriate substrate and stopped. The reaction is then measured visually by the appearance of a colored spot, *e.g.*, by colorimetry.

Under these experimental conditions, a positive reaction is associated with

detection of a colored spot for reactions carried out with a dilution of at least about 1:50, preferably, of at least about 1:500. Therapeutic or prophylactic efficacy of a

polypeptide or polypeptide derivative of the invention can be evaluated as described below.

According to a seventh aspect of the invention, there is provided (i) a composition of matter containing a polypeptide of the invention together with a diluent or carrier; (ii) a pharmaceutical composition containing a therapeutically or prophylactically effective amount of a polypeptide of the invention; (iii) a method for inducing an immune response against *Helicobacter* in a mammal by administering to the mammal an immunogenically effective amount of a polypeptide of the invention to elicit an immune response, *e.g.*, a protective immune response to *Helicobacter*; and (iv) a method for preventing and/or treating a *Helicobacter* (*e.g.*, *H. pylori*, *H. felis*, *H. mustelae*, or *H. heilmanii*) infection, by administering a prophylactic or therapeutic amount of a polypeptide of the invention to an individual in need of such treatment. Additionally, this aspect of the invention includes the use of a polypeptide of the invention in the preparation of a medicament for preventing and/or treating *Helicobacter* infection.

The immunogenic compositions of the invention can be administered by any conventional route in use in the vaccine field, for example, to a mucosal (*e.g.*, ocular, intranasal, pulmonary, oral, gastric, intestinal, rectal, vaginal, or urinary tract) surface or *via* a parenteral (*e.g.*, subcutaneous, intradermal, intramuscular, intravenous, or intraperitoneal) route. The choice of the administration route depends upon a number of parameters, such as the adjuvant used. For example, if a mucosal adjuvant is used, the intranasal or oral route will be preferred, and if a lipid formulation or an aluminum compound is used, a parenteral route will be preferred. In the latter case, the subcutaneous or intramuscular route is most preferred. The choice of administration route can also depend upon the nature of

the vaccine agent. For example, a polypeptide of the invention fused to CTB or to LTB will be best administered to a mucosal surface.

A composition of the invention can contain one or several polypeptides or derivatives of the invention. It can also contain at least one additional

5 *Helicobacter* antigen, such as the urease apoenzyme, or a subunit, fragment, homolog, mutant, or derivative thereof.

For use in a composition of the invention, a polypeptide or polypeptide derivative can be formulated into or with liposomes, such as neutral or anionic liposomes, microspheres, ISCOMS, or virus-like particles (VLPs), to facilitate
10 delivery and/or enhance the immune response. These compounds are readily available to those skilled in the art; for example, see *Liposomes: A Practical Approach (supra)*. Adjuvants other than liposomes can also be used in the invention and are well known in the art (see, for example, the list provided below).

15 Administration can be achieved in a single dose or repeated as necessary at intervals that can be determined by one skilled in the art. For example, a priming dose can be followed by three booster doses at weekly or monthly intervals. An appropriate dose depends on various parameters, including the nature of the recipient (*e.g.*, whether the recipient is an adult or an infant), the
20 particular vaccine antigen, the route and frequency of administration, the presence/absence or type of adjuvant, and the desired effect (*e.g.*, protection and/or treatment), and can be readily determined by one skilled in the art. In general, a vaccine antigen of the invention can be administered mucosally in an amount ranging from about 10 μ g to about 500 mg, preferably from about 1 mg to about

200 mg. For a parenteral route of administration, the dose usually should not exceed about 1 mg, and is, preferably, about 100 μ g.

When used as components of a vaccine, the polynucleotides and polypeptides of the invention can be used sequentially as part of a multi-step immunization process. For example, a mammal can be initially primed with a vaccine vector of the invention, such as a pox virus, *e.g.*, *via* a parenteral route, and then boosted twice with a polypeptide encoded by the vaccine vector, *e.g.*, *via* the mucosal route. In another example, liposomes associated with a polypeptide or polypeptide derivative of the invention can be used for priming, with boosting being carried out mucosally using a soluble polypeptide or polypeptide derivative of the invention, in combination with a mucosal adjuvant (*e.g.*, LT).

Polypeptides and polypeptide derivatives of the invention can also be used as diagnostic reagents for detecting the presence of anti-*Helicobacter* antibodies, *e.g.*, in blood samples. Such polypeptides can be about 5 to about 80, preferably, about 10 to about 50 amino acids in length and can be labeled or unlabeled, depending upon the diagnostic method. Diagnostic methods involving such a reagent are described below.

Upon expression of a polynucleotide molecule of the invention, a polypeptide or polypeptide derivative is produced and can be purified using known methods. For example, the polypeptide or polypeptide derivative can be produced as a fusion protein containing a fused tail that facilitates purification. The fusion product can be used to immunize a small mammal, *e.g.*, a mouse or a rabbit, in order to raise monospecific antibodies against the polypeptide or polypeptide derivative. The eighth aspect of the invention thus provides a monospecific antibody that binds to a polypeptide or polypeptide derivative of the invention.

By "monospecific antibody" is meant an antibody that is capable of reacting with a unique, naturally-occurring *Helicobacter* polypeptide. An antibody of the invention can be polyclonal or monoclonal. Monospecific antibodies can be recombinant, *e.g.*, chimeric (*e.g.*, consisting of a variable region of murine origin and a human constant region), humanized (*e.g.*, a human immunoglobulin constant region and a variable region of animal, *e.g.*, murine, origin), and/or single chain. Both polyclonal and monospecific antibodies can also be in the form of immunoglobulin fragments, *e.g.*, F(ab)'2 or Fab fragments. The antibodies of the invention can be of any isotype, *e.g.*, IgG or IgA, and polyclonal antibodies can be of a single isotype or can contain a mixture of isotypes.

The antibodies of the invention, which can be raised to a polypeptide or polypeptide derivative of the invention, can be produced and identified using standard immunological assays, *e.g.*, Western blot assays, dot blot assays, or ELISA (see, *e.g.*, Coligan *et al.*, *Current Protocols in Immunology*, John Wiley & Sons, Inc., New York, NY, 1994). The antibodies can be used in diagnostic methods to detect the presence of *Helicobacter* antigens in a sample, such as a biological sample. The antibodies can also be used in affinity chromatography methods for purifying a polypeptide or polypeptide derivative of the invention. As is discussed further below, the antibodies can also be used in prophylactic and therapeutic passive immunization methods.

Accordingly, a ninth aspect of the invention provides (i) a reagent for detecting the presence of *Helicobacter* in a biological sample that contains an antibody, polypeptide, or polypeptide derivative of the invention; and (ii) a diagnostic method for detecting the presence of *Helicobacter* in a biological sample, by contacting the biological sample with an antibody, a polypeptide, or a

polypeptide derivative of the invention, so that an immune complex is formed, and detecting the complex as an indication of the presence of *Helicobacter* in the sample or the organism from which the sample was derived. The immune complex is formed between a component of the sample and the antibody,
5 polypeptide, or polypeptide derivative, and that any unbound material can be removed prior to detecting the complex. A polypeptide reagent can be used for detecting the presence of anti-*Helicobacter* antibodies in a sample, *e.g.*, a blood sample, while an antibody of the invention can be used for screening a sample, such as a gastric extract or biopsy sample, for the presence of *Helicobacter*
10 polypeptides.

For use in diagnostic methods, the reagent (*e.g.*, the antibody, polypeptide, or polypeptide derivative of the invention) can be in a free state or can be immobilized on a solid support, such as, for example, on the interior surface of a tube or on the surface, or within pores, of a bead. Immobilization can be
15 achieved using direct or indirect means. Direct means include passive adsorption (*i.e.*, non-covalent binding) or covalent binding between the support and the reagent. By "indirect means" is meant that an anti-reagent compound that interacts with the reagent is first attached to the solid support. For example, if a polypeptide reagent is used, an antibody that binds to it can serve as an anti-reagent, provided
20 that it binds to an epitope that is not involved in recognition of antibodies in biological samples. Indirect means can also employ a ligand-receptor system, for example, a molecule, such as a vitamin, can be grafted onto the polypeptide reagent and the corresponding receptor can be immobilized on the solid phase. This concept is illustrated by the well known biotin-streptavidin system.
25 Alternatively, indirect means can be used, *e.g.*, by adding to the reagent a peptide

tail, chemically or by genetic engineering, and immobilizing the grafted or fused product by passive adsorption or covalent linkage of the peptide tail.

According to a tenth aspect of the invention, there is provided a process for purifying, from a biological sample, a polypeptide or polypeptide derivative of the invention, which involves carrying out antibody-based affinity chromatography with the biological sample, wherein the antibody is a monospecific antibody of the invention.

For use in a purification process of the invention, the antibody can be polyclonal or monospecific, and preferably is of the IgG type. Purified IgGs can be prepared from an antiserum using standard methods (see, *e.g.*, Coligan *et al.*, *supra*). Conventional chromatography supports, as well as standard methods for grafting antibodies, are described, for example, by Harlow *et al.* (*Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1988).

Briefly, a biological sample, such as an *H. pylori* extract, preferably in a buffer solution, is applied to a chromatography material, which is, preferably, equilibrated with the buffer used to dilute the biological sample, so that the polypeptide or polypeptide derivative of the invention (*i.e.*, the antigen) is allowed to adsorb onto the material. The chromatography material, such as a gel or a resin coupled to an antibody of the invention, can be in batch form or in a column. The unbound components are washed off and the antigen is eluted with an appropriate elution buffer, such as a glycine buffer, a buffer containing a chaotropic agent, *e.g.*, guanidine HCl, or a buffer having high salt concentration (*e.g.*, 3 M MgCl₂). Eluted fractions are recovered and the presence of the antigen is detected, *e.g.*, by measuring the absorbance at 280 nm.

An antibody of the invention can be screened for therapeutic efficacy as follows. According to an eleventh aspect of the invention, there is provided (i) a composition of matter containing a monospecific antibody of the invention, together with a diluent or carrier; (ii) a pharmaceutical composition containing a therapeutically or prophylactically effective amount of a monospecific antibody of the invention, and (iii) a method for treating or preventing *Helicobacter* (e.g., *H. pylori*, *H. felis*, *H. mustelae*, or *H. heilmanii*) infection, by administering a therapeutic or prophylactic amount of a monospecific antibody of the invention to an individual in need of such treatment. In addition, the eleventh aspect of the invention includes the use of a monospecific antibody of the invention in the preparation of a medicament for treating or preventing *Helicobacter* infection.

The monospecific antibody can be polyclonal or monoclonal, and is, preferably, predominantly of the IgA isotype. In passive immunization methods, the antibody is administered to a mucosal surface of a mammal, e.g., the gastric mucosa, e.g., orally or intragastrically, optionally, in the presence of a bicarbonate buffer. Alternatively, systemic administration, not requiring a bicarbonate buffer, can be carried out. A monospecific antibody of the invention can be administered as a single active agent or as a mixture with at least one additional monospecific antibody specific for a different *Helicobacter* polypeptide. The amount of antibody and the particular regimen used can be readily determined by one skilled in the art. For example, daily administration of about 100 to 1,000 mg of antibody over one week, or three doses per day of about 100 to 1,000 mg of antibody over two or three days, can be effective regimens for most purposes.

Therapeutic or prophylactic efficacy can be evaluated using standard methods in the art, e.g., by measuring induction of a mucosal immune response or

induction of protective and/or therapeutic immunity, using, *e.g.*, the *H. felis* mouse model and the procedures described by Lee *et al.* (Eur. J. Gastroenterology & Hepatology 7:303, 1995) or Lee *et al.* (J. Infect. Dis. 172:161, 1995). Those skilled in the art will recognize that the *H. felis* strain of the model can be replaced
5 with another *Helicobacter* strain. For example, the efficacy of polynucleotide molecules and polypeptides from *H. pylori* is, preferably, evaluated in a mouse model using an *H. pylori* strain. Protection can be determined by comparing the degree of *Helicobacter* infection in the gastric tissue assessed by, for example, urease activity, bacterial counts, or gastritis, to that of a control group. Protection
10 is shown when infection is reduced by comparison to the control group. Such an evaluation can be made for polynucleotides, vaccine vectors, polypeptides, and polypeptide derivatives, as well as for antibodies of the invention.

For example, various doses of an antibody of the invention can be administered to the gastric mucosa of mice previously challenged with an *H. pylori*
15 strain, as described, *e.g.*, by Lee *et al.* (*supra*). Then, after an appropriate period of time, the bacterial load of the mucosa can be estimated by assessing urease activity, as compared to a control. Reduced urease activity indicates that the antibody is therapeutically effective.

Adjuvants that can be used in any of the vaccine compositions described
20 above are described as follows. Adjuvants for parenteral administration include, for example, aluminum compounds, such as aluminum hydroxide, aluminum phosphate, and aluminum hydroxy phosphate. The antigen can be precipitated with, or adsorbed onto, the aluminum compound using standard methods. Other adjuvants, such as RIBI (ImmunoChem, Hamilton, MT), can also be used in
25 parenteral administration.

Adjuvants that can be used for mucosal administration include, for example, bacterial toxins, *e.g.*, the cholera toxin (CT), the *E. coli* heat-labile toxin (LT), the *Clostridium difficile* toxin A, the *pertussis* toxin (PT), and combinations, subunits, toxoids, or mutants thereof. For example, a purified preparation of native cholera toxin subunit B (CTB) can be used. Fragments, homologs, derivatives, and fusions to any of these toxins can also be used, provided that they retain adjuvant activity. Preferably, a mutant having reduced toxicity is used. Suitable mutants are described, *e.g.*, in WO 95/17211 (Arg-7-Lys CT mutant), WO 96/6627 (Arg-192-Gly LT mutant), and WO 95/34323 (Arg-9-Lys and Glu-129-Gly PT mutant). Additional LT mutants that can be used in the methods and compositions of the invention include, *e.g.*, Ser-63-Lys, Ala-69-Gly, Glu-110-Asp, and Glu-112-Asp mutants. Other adjuvants, such as the bacterial monophosphoryl lipid A (MPLA) of, *e.g.*, *E. coli*, *Salmonella minnesota*, *Salmonella typhimurium*, or *Shigella flexneri*; saponins, and polylactide glycolide (PLGA) microspheres, can also be used in mucosal administration. Adjuvants useful for both mucosal and parenteral administrations, such as polyphosphazene (WO 95/2415), can also be used.

Any pharmaceutical composition of the invention, containing a polynucleotide, polypeptide, polypeptide derivative, or antibody of the invention, can be manufactured using standard methods. It can be formulated with a pharmaceutically acceptable diluent or carrier, *e.g.*, water or a saline solution, such as phosphate buffer saline, optionally, including a bicarbonate salt, such as sodium bicarbonate, *e.g.*, 0.1 to 0.5 M. Bicarbonate can advantageously be added to compositions intended for oral or intragastric administration. In general, a diluent or carrier can be selected on the basis of the mode and route of administration, and

standard pharmaceutical practice. Suitable pharmaceutical carriers and diluents, as well as pharmaceutical necessities for their use in pharmaceutical formulations, are described in *Remington's Pharmaceutical Sciences*, a standard reference text in this field and in the USP/NF.

5 The invention also includes methods in which gastroduodenal infections, such as *Helicobacter* infection, are treated by oral administration of a *Helicobacter* polypeptide of the invention and a mucosal adjuvant, in combination with an antibiotic, an antisecretory agent, a bismuth salt, an antacid, sucralfate, or a combination thereof. Examples of such compounds that can be administered with
10 the vaccine antigen and an adjuvant are antibiotics, including, *e.g.*, macrolides, tetracyclines, β -lactams, aminoglycosides, quinolones, penicillins, and derivatives thereof (specific examples of antibiotics that can be used in the invention include, *e.g.*, amoxicillin, clarithromycin, tetracycline, metronidazole, erythromycin, cefuroxime, and erythromycin); antisecretory agents, including, *e.g.*, H₂-receptor
15 antagonists (*e.g.*, cimetidine, ranitidine, famotidine, nizatidine, and roxatidine), proton pump inhibitors (*e.g.*, omeprazole, lansoprazole, and pantoprazole), prostaglandin analogs (*e.g.*, misoprostil and enprostil), and anticholinergic agents (*e.g.*, pirenzepine, telenzepine, carbenoxolone, and proglumide); and bismuth salts, including colloidal bismuth subcitrate, tripotassium dicitrate bismuthate, bismuth
20 subsalicylate, bicitropeptide, and pepto-bismol (see, *e.g.*, Goodwin *et al.*, *Helicobacter pylori*, *Biology and Clinical Practice*, CRC Press, Boca Raton, FL, pp 366-395, 1993; Physicians' Desk Reference, 49th edn., Medical Economics Data Production Company, Montvale, New Jersey, 1995). In addition, compounds containing more than one of the above-listed components coupled together, *e.g.*,
25 ranitidine coupled to bismuth subcitrate, can be used. The invention also includes

compositions for carrying out these methods, *i.e.*, compositions containing a *Helicobacter* antigen (or antigens) of the invention, an adjuvant, and one or more of the above-listed compounds, in a pharmaceutically acceptable carrier or diluent.

Amounts of the above-listed compounds used in the methods and compositions of the invention can readily be determined by one skilled in the art. In addition, one skilled in the art can readily design treatment/immunization schedules. For example, the non-vaccine components can be administered on days 1-14, and the vaccine antigen + adjuvant can be administered on days 7, 14, 21, and 28.

Methods and pharmaceutical compositions of the invention can be used to treat or to prevent *Helicobacter* infections and, accordingly, gastroduodenal diseases associated with these infections, including acute, chronic, and atrophic gastritis, and peptic ulcer diseases, *e.g.*, gastric and duodenal ulcers.

All of the clones of the invention were originally isolated by a transposon shuttle mutagenesis method. Briefly, in this method, a TnMax9 mini-*blaM* transposon was used for insertional mutagenesis of an *H. pylori* gene library established in *E. coli*. 192 *E. coli* clones expressing active β -lactamase fusion proteins were obtained, indicating that the corresponding target plasmids carry *H. pylori* genes encoding extracytoplasmic proteins. Individual mutants were transferred onto the chromosome of *H. pylori* P1 or P12 by natural transformation, resulting in 135 distinct *H. pylori* mutants. This method is described in further detail, as follows.

The transposon TnMax9 (Kahrs *et al.*, Gene 167:53, 1995) was used to generate mutations in an *H. pylori* library in *E. coli*. As illustrated in Fig. 1A, TnMax9 contains, in addition to a cat_{GC} -resistance gene close to the inverted repeat

(IR), an unexpressed open reading frame encoding β -lactamase without a promoter or leader sequence (mature β -lactamase, *blaM*; Kahrs *et al.*, *supra*). For production of extracytoplasmic BlaM fusion proteins resulting in ampicillin-resistant (amp^R) clones, expression of the cloned *H. pylori* genes in *E. coli* is
5 obligatory. The minimal vector pMin2 (Kahrs *et al.*, *supra*; see Fig. 1B), containing a weak constitutive promoter (P_{iga}) upstream of the multiple cloning site, was used for construction of the *H. pylori* library to ensure expression of *H. pylori* genes in *E. coli*.

In construction of the library, *H. pylori* DNA was partially digested with
10 *Sau3A* and *HpaII*, size fractionated by preparative agarose gel electrophoresis, and 3-6 kb fragments were ligated into the *BglIII* and *ClaI* sites of pMin2. The library was introduced into *E. coli* strain E181(pTnMax9), which is a derivative of HB101 containing the TnMax9 transposon, by electroporation. This generated approximately 2,400 independent transformants. More than 95% of the plasmids
15 contained an insert of between 3 and 6 kb, showing that the 1.7 Mb *H. pylori* chromosome was statistically covered. Since not every plasmid could be expected to contain a target gene carrying an export signal, the library was partitioned into a total of 198 pools (24 pools of 20 clones and 174 pools of 11 clones). Using a cotton swab, either eleven or twenty individual colonies were inoculated in 0.5 mL
20 LB medium in a eppendorf tubes, vortexed, and 100 mL of the suspension was spread on LB agar plates supplemented with tetracycline and chloramphenicol to select for maintenance of both plasmids. Insertion of TnMax9 into the target plasmids was induced with 100 mM isopropyl-b-D-thiogalactoside (IPTG) separately for each pool (Haas *et al.*, Gene 130:23-21, 1993). Plasmids were
25 transferred into E145 by triparental mating, in which 25 mL of the donor strain

(E181), 25 mL of the mobilisator (HB101(pRK2013)), and 50 mL of the recipient strain (E145) were mixed from corresponding bacterial suspensions ($\text{O.D.}_{550} = 10$). The matings were performed for 2-3 hours at 37°C on nitrocellulose filters, which were placed on LB plates. Bacteria were suspended in 1 mL LB and aliquots were spread on LB plates containing chloramphenicol, tetracycline, and rifampicin. Each pool gave rise to chloramphenicol-resistant transconjugates in E145, demonstrating that both transposition and conjugation were successful. Generally, several thousand chloramphenicol-resistant transconjugates were obtained, but the number of amp^R colonies varied in different pools, ranging from one to several hundred colonies. Two amp^R colonies from each positive pool were isolated, plasmid DNA was extracted, and the DNA was characterized by further restriction analysis. Only those *TnMax9* insertions of a single pool that mapped in obviously different plasmid clones, or in markedly different regions of the same clone, were used further.

From 158 of the 198 pools, ampicillin-resistant E145 transconjugates were obtained (80%), showing that in several pools, *TnMax9* inserted into expressed genes, resulting in production of extracytoplasmic BlaM fusion proteins. Thus, a total of 192 amp^R E145 clones could be isolated by conjugal transfer of plasmids from 198 pools.

To analyze the mutant library, it was determined whether defined gene sequences inactivated by *TnMax9* were represented once or several times in the whole library. Five transposon-containing plasmids conferring an amp^R phenotype to E145 (pMu7, pMu13, pMu75, pMu94, and pMu110) were randomly selected and DNA fragments flanking the *TnMax9* insert were isolated and used as probes in Southern hybridization of 120 amp^R clones. The hybridization probes isolated

from clones pMu7, pMu75, and pMu94 were between 0.9 and 1.1 kb in size, and hybridized exclusively with the inserts of the homologous plasmids. In contrast, the *TnMax9* flanking regions of clones pMu13 and pMu110 were 4.0 kb and 5.5 kb, respectively. They each hybridized with the homologous plasmids, and with one additional clone of the library. Such a result was expected, since the chance of a probe to find a homologous sequence in the library should be higher, the longer the hybridization probes.

In order to verify the insertion of the transposon into distinct ORFs encoding putative exported proteins, the *TnMax9*-flanking DNA of five representative *amp^R* mutant clones (pMu7, pMu12, pMu18, pMu20, and pMu26) was sequenced, taking advantage of the M13 forward and reverse primers on *TnMax9* (Fig. 1A). This analysis revealed that the mini-transposon was inserted into different sequences in each plasmid, thereby interrupting ORFs encoding putative proteins. For two clones, the sequences located upstream of the *blaM* gene revealed a putative ribosome-binding site and a potential translational start codon (ATG). Other clones either revealed an ORF spanning the complete sequence (approximately 400 base pairs upstream and downstream of the *TnMax9* insertion) or terminating shortly after the site of *TnMax9* insertion. The partial protein sequences from different ORFs were used for database searches, but no significant homologies with known proteins were found.

In a further approach, it was determined whether a known gene, like *vacA*, encoding the extracellular vacuolating cytotoxin of *H. pylori*, could be identified using this method and how often such a mutation would be represented in the mutant library. A total cell lysates of the 135 mutants were tested in an immunoblot using the *H. pylori* cytotoxin-specific rabbit antiserum AK197

(Schmitt *et al.*, Mol. Microbiol. 12:307-319, 1994). Two mutants were identified that no longer produced the cytotoxin antigen (mutants P1-26 and P1-47) and partial DNA sequencing of the insertion sites revealed that TnMax9 was inserted at distinct positions in the *vacA* gene, 56 and 53 codons downstream of the ATG start codon, respectively.

Thus, the characterization of the mutant collection confirmed that a representative gene library was constructed in *E. coli*, in which target genes encoding exported *H. pylori* proteins were efficiently tagged by TnMax9.

In order to establish a collection of mutants lacking distinct exported proteins, the mutations had to be transferred back into the *H. pylori* chromosome. By means of natural transformation, 86 plasmids could be transformed into the original strain P1. *H. pylori* strains P1 or P12, which were naturally competent for DNA transformation, were transformed with circular plasmid DNA (0.2-0.5 mg/transformation). Transformations to streptomycin resistance were performed with chromosomal DNA (1 mg/transformation), isolated from a streptomycin-resistant NCTC11637 *H. pylori* mutant according to the procedure described in Haas *et al.* (Mol. Microbiol. 8:753-760). Selection was performed on serum plates containing 4 mg/mL chloramphenicol or 500 mg/mL streptomycin. The transformation frequency for a given mutant was calculated as the number of chloramphenicol-, streptomycin-, or erythromycin-resistant colonies per cfu (average of three experiments). The *blaM* gene was deleted by *NotI* digestion, and the plasmid religated, in those plasmids that did not transform strain P1 directly. This procedure, which resulted in a twenty- to thirty-fold higher frequency of transformation, as compared to the same plasmid containing *blaM*, resulted in 36 additional mutants strain P1. The *blaM*-deletion plasmids that still did not

transform strain P1 were used to transform the heterologous *H. pylori* strain P12, possessing an approximately 10-fold higher transformation frequency compared to P1. This resulted in thirteen further mutants.

Thus, from the 192 amp^R plasmids a total of 135 *H. pylori* mutants (122 mutants in P1 and 13 mutants in P12) were finally obtained by selection for chloramphenicol resistance (70%). The transformation frequency varied between different plasmids in the range of 1×10^{-5} - 1×10^{-7} . The remaining plasmids did not result in any transformants. The collection was frozen as individual mutants in stock cultures at -70°C. To verify the correct insertion of the mini-transposon into the *H. pylori* chromosome, ten representative mutants were tested by Southern hybridization of chromosomal DNA using *cat*_{GC} DNA and the vector pMin2 as probes. Consistent with our previous experience concerning TnMax9-based shuttle mutagenesis of *H. pylori*, the mini-transposon was, in all cases, inserted into the chromosome without integration of the vector DNA, which probably means by a double cross-over, rather than by a single cross-over event. As judged from the hybridization pattern obtained with the *cat* gene as a probe, it appears that TnMax9 is located in different regions of the chromosome, showing that distinct target genes have been interrupted in individual mutants.

The mutants were analyzed for motility, transformation competence, and adherence to KatoIII cells. Screening of the *H. pylori* mutant collection allowed identification of mutants impaired in motility, natural transformation competence, and adherence to gastric epithelial cell lines. Motility mutants could be grouped into distinct classes: (i) mutants lacking the major flagellin subunit FlaA and intact flagella; (ii) mutants with apparently normal flagella, but reduced motility; and (iii) mutants with obviously normal flagella, but completely

abolished motility. Two independent mutations, which exhibited defects in natural competence for genetic transformation, mapped to different genetic loci. In addition, two independent mutants were isolated by their failure to bind to the human gastric carcinoma cell line KatoIII. Both mutants carried a transposon in the same gene, approximately 0.8 kb apart, and showed decrease autoagglutination, when compared to the wild type strain.

Sequences of clones obtained using the above-described transposon shuttle mutagenesis method were used to identify intact genes, lacking inserted transposons, in the *H. pylori* genome, as is described below in Example 1.

The invention is further illustrated by the following examples. Example 1 describes identification of genes, such as genes that encode the polypeptides of the invention, in the *Helicobacter* genome, as well as identification of leader sequences, and primer design for amplification of genes lacking signal sequences. Example 2 describes cloning of DNA encoding GHPO 136, GHPO 191, GHPO 411, GHPO 419, GHPO 724, and GHPO 427 into a vector that provides a histidine tag, and production and purification of the resulting his-tagged fusion proteins. Example 3 describes methods for cloning DNA encoding the polypeptides of the invention so that they can be produced without his-tags, and Example 4 describes methods for purifying recombinantly produced polypeptides of the invention.

EXAMPLE 1: Identification of genes in the *H. pylori* genome, identification of leader sequences, and primer design for amplification of genes lacking signal sequences

1.A. Creating *H. pylori* genomic databases

5 The *H. pylori* genome was provided as a text file containing a single contiguous string of nucleotides that had been determined to be 1.76 Megabases in length. The complete genome was split into 17 separate files using the program SPLIT (Creativity in Action), giving rise to 16 contigs, each containing 100,000 nucleotides, and a 17th contig containing the remaining 76,000 nucleotides. A
10 header was added to each of the 17 files using the format: >hpg0.txt (representing contig 1), .hpg1.txt (representing contig 2), etc. The resulting 17 files, named hpg0 through hpg16, were then copied together to form one file that represented the plus strand of the complete *H. pylori* genome. The constructed database was given the designation "H." A negative strand database of the *H. pylori* genome
15 was created similarly by first creating a reverse complement of the positive strand using the program SeqPup (D.G. Gilbert, Indiana University Biology Department) and then performing the same procedure as described above for the plus strand. This database was given the designation "N."

20 The regions predicted to encode open reading frames (ORFs) were defined for the complete *H. pylori* genome using the program GENEMARK™ (Borodovsky *et al.*, Comp. Chem. 17:123, 1993). A database was created from a text file containing an annotated version of all ORFs predicted to be encoded by the *H. pylori* genome for both the plus and minus strands, and was given the designation "O." Each ORF was assigned a number indicating its location on the

genome and its position relative to other genes. No manipulation of the text file was required.

1.B. Searching the *H. pylori* databases

The databases constructed as is described above were searched using the
5 program FASTA (Pearson *et al.*, Proc. Natl. Acad. Sci. USA 85:2444-2448, 1988).
FASTA was used for searching either a DNA sequence against either of the gene
databases (“H” and/or “N”), or a peptide sequence against the ORF library (“O”).
TFASTX was used to search a peptide sequence against all possible reading
frames of a DNA database (“H” and/or “N” libraries). Potential frameshifts also
10 being resolved, FASTX was used for searching the translated reading frames of a
DNA sequence against either a DNA database, or a peptide sequence against the
protein database.

1.C. Isolation of DNA sequences from the *H. pylori* genome

The FASTA searches against the constructed DNA databases identified
15 exact nucleotide coordinates on one or more of the isolated contigs, and therefore
the location of the target DNA. Once the exact location of the target sequence was
known, the contig identified to carry the gene was exported into the software
package MapDraw (DNASar, Inc.) and the gene was isolated. Gene sequences
with flanking DNA was then excised and copied into the EditSeq. Software
20 package (DNASar, Inc.) for further analysis.

1.D. Identification of leader sequences

The deduced protein encoded by a target gene sequence is analyzed using the PROTEAN software package (DNASTar, Inc.). This analysis predicts those areas of the protein that are hydrophobic by using the Kyte-Doolittle algorithm, and identifies any potential polar residues preceding the hydrophobic core region, which is typical for many leader sequences. For confirmation, the target protein is then searched against a PROSITE database (DNASTar, Inc.) consisting of motifs and signatures. Characteristic of many leader sequences and hydrophobic regions in general, is the identification of predicted prokaryotic lipid attachment sites. Where confirmation between the two approaches is apparent at the N-terminus of any protein, putative cleavage sites are sought. Specifically, this includes the presence of either an Alanine (A), Serine (S), or Glycine (G) residue immediately after the core hydrophobic region. In the case of lipoproteins, a Cysteine (C) residue would be identified as the +1 residue, post-cleavage.

1.E. Rational design of PCR primers based on the identification of leader sequences

In order to clone gene sequences as N-terminus translational fusions for the generation of recombinant proteins with N-terminal Histidine tags, the gene sequence that specifies the leader sequence is omitted. The 5'-end of the gene-specific portion of the N-terminal primer is designed to start at the first codon beyond the cleavage site. In the case of lipoproteins, the 5'- end of the N-terminal primer begins at the second codon, immediately after the modifiable residue at position +1 post-cleavage. The omission of the leader sequence from the recombinant allows for one-step purification, and potential problems associated

with insertion of leader sequences in the membrane of the host strain carrying the hybrid construct are avoided.

EXAMPLE 2: Preparation of isolated DNA encoding GHPO 136, GHPO 191, GHPO 411, GHPO 419, GHPO 724, and GHPO 427, and production of these polypeptides as histidine-tagged fusion proteins

2.A. Preparation of genomic DNA from *Helicobacter pylori*

Helicobacter pylori strain ORV2001, stored in LB medium containing 50% glycerol at -70°C, is grown on Colombia agar containing 7% sheep blood for 48 hours under microaerophilic conditions (8-10% CO₂, 5-7% O₂, 85-87% N₂).

Cells are harvested, washed with phosphate buffer saline (PBS) (pH 7.2), and DNA is then extracted from the cells using the Rapid Prep Genomic DNA Isolation kit (Pharmacia Biotech).

2.B. PCR amplification

DNA molecules encoding GHPO 136, GHPO 191, GHPO 408, GHPO 411, GHPO 419, GHPO 724, and GHPO 427 are amplified from genomic DNA, as can be prepared as is described above, by the Polymerase Chain Reaction (PCR) using the following primers:

GHPO 136:

N-terminal primer:

5'-CGCGGATCCGAAATAGGGTTGTTTTTAATTTTC-3' (SEQ ID NO:101);

and

C-terminal primer:

5'-CCGCTCGAGTTAAAAAAGAGTTTGTATAA-3' (SEQ ID NO:102).

GHPO 191:

N-terminal primer:

5'-GGGGATCCTTGGTAGAATTGAATCA-3' (SEQ ID NO:103); and

5 C-terminal primer:

5'-GGAATTCCTAAAACAAGAACGCG-3' (SEQ ID NO:104).

GHPO 411:

N-terminal primer:

5'-GGGGATCCTTTTTTCAAAAACAATA-3' (SEQ ID NO:105); and

10 C-terminal primer:

5'-GGAATTCTCACATTGTTTGTCTC-3' (SEQ ID NO:106).

GHPO 419:

N-terminal primer:

5'-GCGGATCCCAATTTCAAAAAGCC-3' (SEQ ID NO:107); and

15 C-terminal primer:

5'-CCGCTCGAACTAAAACTATAAACG-3' (SEQ ID NO:108).

GHPO 724:

N-terminal primer:

5'-CGCGGATCCGAGATTTTGAAAGGTTGGTAATG-3' (SEQ ID NO:109);

20 and

C-terminal primer:

5'-CCGCTCGAGCTACATCCTTTTACTATAACC-3' (SEQ ID NO:110).

GHPO 427:

N-terminal primer:

25 5'-GCGGATCCGGGTATTATTCAGAAG-3' (SEQ ID NO:111); and

C-terminal primer:

5'-CCGCTCGAGTTAAAATTTGCTCGC-3' (SEQ ID NO:112).

The N-terminal and C-terminal primers for each clone both include a 5' clamp and a restriction enzyme recognition sequence for cloning purposes (*Bam*HI (GGATCC) and *Xho*I (CTCGAG) recognition sequences).

Amplification of gene-specific DNA is carried out using Vent DNA Polymerase (New England Biolabs) or Taq DNA polymerase (Appligene), according to the manufacturer's instructions. The reaction mixture, which is brought to a final volume of 100 μ L with distilled water, is as follows:

10	dNTPs mix	200 μ M
	10x ThermoPol buffer	10 μ L
	primers	300 nM each
	DNA template	50 ng
15	Heat-stable DNA polymerase	2 units

Appropriate amplification reaction conditions can readily be determined by one skilled in the art. In the present case, Vent DNA polymerase (New England Biolabs) was used to amplify GHPO 136, GHPO 191, GHPO 411, GHPO 419, GHPO 724, and GHPO 427 as follows. For GHPO 136, a denaturing step was carried out at 97°C for 30 seconds, followed by an annealing step at 55°C for 45 seconds, and an extension step at 72°C for 1 minute and 30 seconds. Twenty five cycles were carried out. For GHPO 191 and GHPO 427, an initial denaturing step was carried out at 94°C for 5 minutes, and was followed by a number of cycles (20 for GHPO 191 and 25 for GHPO 427), including a denaturing step at 94°C for 30 seconds, an annealing step at 50°C for 30 seconds, and an extension

step at 72°C for thirty seconds. The 20 cycles were followed by a final elongation step at 72°C for 7 minutes. For GHPO 411, an initial denaturing step was carried out at 94°C for 5 minutes, and was followed by 25 cycles, including a denaturing step at 94°C for 30 seconds, an annealing step at 50°C for 30 seconds, and an extension step at 72°C for 30 seconds. The 25 cycles were followed by a final elongation step at 72°C for 7 minutes. For GHPO 419 the same reaction conditions were used as for GHPO 411, except that 30 cycles were carried out for GHPO 419, instead of 25. For GHPO 724, twenty five cycles, including a denaturing step at 97°C for 30 seconds, an annealing step at 55°C for 1 minute, and an elongation step at 72°C for 7 minutes, were carried out.

2.C. Transformation and selection of transformants

A single PCR product is thus amplified and is then digested at 37°C for 2 hours with *Bam*HI and *Xho*I concurrently in a 20 µL reaction volume. The digested product is ligated to similarly cleaved pET28a (Novagen) that is dephosphorylated prior to the ligation by treatment with Calf Intestinal Alkaline Phosphatase (CIP). The gene fusion constructed in this manner allows one-step affinity purification of the resulting fusion protein because of the presence of histidine residues at the N-terminus of the fusion protein, which are encoded by the vector.

The ligation reaction (20 µL) is carried out at 14°C overnight and then is used to transform 100 µL fresh *E. coli* XL1-blue competent cells (Novagen). The cells are incubated on ice for 2 hours, heat-shocked at 42°C for 30 seconds, and returned to ice for 90 seconds. The samples are then added to 1 mL LB broth in the absence of selection and grown at 37°C for 2 hours. The cells are plated out

on LB agar containing kanamycin (50 $\mu\text{g/mL}$) at a 10x and neat dilution and incubated overnight at 37°C. The following day, 50 colonies are picked onto secondary plates and incubated at 37°C overnight.

Five colonies are picked into 3 mL LB broth supplemented with
5 kanamycin (100 $\mu\text{g/mL}$) and are grown overnight at 37°C. Plasmid DNA is extracted using the Quiagen mini-prep. method and is quantitated by agarose gel electrophoresis.

PCR is performed with the gene-specific primers under the conditions set forth above and transformant DNA is confirmed to contain the desired insert.

10 If PCR-positive, one of the five plasmid DNA samples (500 ng) extracted from the *E. coli* XL1-blue cells is used to transform competent BL21 (λDE3) *E. coli* competent cells (Novagen; as described previously). Transformants (10) are picked onto selective kanamycin (50 $\mu\text{g/mL}$) containing LB agar plates and stored as a research stock in LB containing 50% glycerol.

15 **2.D. Purification of recombinant proteins**

One mL of frozen glycerol stock prepared as described in 2.C. is used to inoculate 50 mL of LB medium containing 25 $\mu\text{g/mL}$ of kanamycin in a 250 mL Erlenmeyer flask. The flask is incubated at 37°C for 2 hours or until the absorbance at 600 nm (OD_{600}) reaches 0.4-1.0. The culture is stopped from
20 growing by placing the flask at 4°C overnight. The following day, 10 mL of the overnight culture are used to inoculate 240 mL LB medium containing kanamycin (25 $\mu\text{g/mL}$), with the initial OD_{600} about 0.02-0.04. Four flasks are inoculated for each ORF. The cells are grown to an OD_{600} of 1.0 (about 2 hours at 37°C), a 1 mL sample is harvested by centrifugation, and the sample is analyzed by SDS-

PAGE to detect any leaky expression. The remaining culture is induced with 1 mM IPTG and the induced cultures are grown for an additional 2 hours at 37°C.

The final OD₆₀₀ is taken and the cells are harvested by centrifugation at 5,000 x g for 15 minutes at 4°C. The supernatant is discarded and the pellets are resuspended in 50 mM Tris-HCl (pH 8.0), 2 mM EDTA. Two hundred and fifty mL of buffer are used for 1 L of culture and the cells are recovered by centrifugation at 12,000 x g for 20 minutes. The supernatant is discarded and the pellets are stored at -45°C.

2. E. Protein purification

Pellets obtained from 2.D. are thawed and resuspended in 95 mL of 50 mM Tris-HCl (pH 8.0). Pefabloc and lysozyme are added to final concentrations of 100 µM and 100 µg/mL, respectively. The mixture is homogenized with magnetic stirring at 5°C for 30 minutes. Benzonase (Merck) is added at a 1 U/mL final concentration, in the presence of 10 mM MgCl₂, to ensure total digestion of the DNA. The suspension is sonicated (Branson Sonifier 450) for 3 cycles of 2 minutes each at maximum output. The homogenate is centrifuged at 19,000 x g for 15 minutes and both the supernatant and the pellet are analyzed by SDS-PAGE to detect the cellular location of the target protein in the soluble or insoluble fractions, as is described further below.

2.E.1. Soluble fraction

If the target protein is produced in a soluble form (*i.e.*, in the supernatant obtained in 2.E.) NaCl and imidazole are added to the supernatant to final concentrations of 50 mM Tris-HCl (pH 8.0), 0.5 M NaCl, and 10 mM imidazole (buffer A). The mixture is filtered through a 0.45 µm membrane and loaded onto

an IMAC column (Pharmacia HiTrap chelating Sepharose; 1 mL), which has been charged with nickel ions according to the manufacturer's recommendations. After loading, the column is washed with 50 column volumes of buffer A and the recombinant target protein is eluted with 5 mL of buffer B (50 mM Tris-HCl (pH 8.0), 0.5 M NaCl, 500 mM imidazole).

The elution profile is monitored by measuring the absorbance of the fractions at 280 nm. Fractions corresponding to the protein peak are pooled, dialyzed against PBS containing 0.5 M arginine, filtered through a 0.22 µm membrane, and stored at -45°C.

2.E.2. Insoluble fraction

If the target protein is expressed in the insoluble fraction (pellets obtained from 2.E.), purification is conducted under denaturing conditions. NaCl, imidazole, and urea are added to the resuspended pellet to final concentrations of 50 mM Tris-HCl (pH 8.0), 0.5 M NaCl, 10 mM imidazole, and 6 M urea (buffer C). After complete solubilization, the mixture is filtered through a 0.45 µm membrane and loaded onto an IMAC column.

The purification procedures on the IMAC column are the same as described in 2.E.1., except that 6 M urea is included in all buffers used and 10 column volumes of buffer C are used to wash the column after protein loading, instead of 50 column volumes.

The protein fractions eluted from the IMAC column with buffer D (buffer C containing 500 mM imidazole) are pooled. Arginine is added to the solution to final concentration of 0.5 M and the mixture is dialyzed against PBS containing 0.5 M arginine and various concentrations of urea (4 M, 3 M, 2 M, 1 M, and 0.5

M) to progressively decrease the concentration of urea. The final dialysate is filtered through a 0.22 μ m membrane and stored at -45°C.

Alternatively, when the above purification process is not as efficient as it should be, two other processes may be used as follows. A first alternative involves the use of a mild denaturant, N-octyl glucoside (NOG). Briefly, a pellet obtained in 2.E. is homogenized in 5 mM imidazole, 500 mM sodium chloride, 20 mM Tris-HCl (pH 7.9) by microfluidization at a pressure of 15,000 psi and is clarified by centrifugation at 4,000-5,000 x g. The pellet is recovered, resuspended in 50 mM NaPO₄ (pH 7.5) containing 1-2 % weight /volume NOG, and homogenized. The NOG-soluble impurities are removed by centrifugation. The pellet is extracted once more by repeating the preceding extraction step. The pellet is dissolved in 8 M urea, 50 mM Tris (pH 8.0). The urea-solubilized protein is diluted with an equal volume of 2 M arginine, 50 mM Tris (pH 8.0), and is dialyzed against 1 M arginine for 24-48 hours to remove the urea. The final dialysate is filtered through a 0.22 μ m membrane and stored at -45°C.

A second alternative involves the use of a strong denaturant, such as guanidine hydrochloride. Briefly, a pellet obtained in 2.E. is homogenized in 5 mM imidazole, 500 mM sodium chloride, 20 mM Tris-HCl (pH 7.9) by microfluidization at a pressure of 15,000 psi and clarified by centrifugation at 4,000-5,000 x g. The pellet is recovered, resuspended in 6 M guanidine hydrochloride, and passed through an IMAC column charged with Ni⁺⁺. The bound antigen is eluted with 8 M urea (pH 8.5). Beta-mercaptoethanol is added to the eluted protein to a final concentration of 1 mM, then the eluted protein is passed through a Sephadex G-25 column equilibrated in 0.1 M acetic acid. Protein

eluted from the column is slowly added to 4 volumes of 50 mM phosphate buffer (pH 7.0). The protein remains in solution.

2.F. Evaluation of the protective activity of the purified protein

Groups of 10 OF1 mice (IFFA Credo) are immunized rectally with 25 μ g of the purified recombinant protein, admixed with 1 μ g of cholera toxin (Berna) in physiological buffer. Mice are immunized on days 0, 7, 14, and 21. Fourteen days after the last immunization, the mice are challenged with *H. pylori* strain ORV2001 grown in liquid media (the cells are grown on agar plates, as described in 2.A., and, after harvest, the cells are resuspended in Brucella broth; the flasks are then incubated overnight at 37°C). Fourteen days after challenge, the mice are sacrificed and their stomachs are removed. The amount of *H. pylori* is determined by measuring the urease activity in the stomach and by culture.

2.G. Production of monospecific polyclonal antibodies

2.G.1. Hyperimmune rabbit antiserum

New Zealand rabbits are injected both subcutaneously and intramuscularly with 100 μ g of a purified fusion polypeptide, as obtained in 2.E.1. or 2.E.2., in the presence of Freund's complete adjuvant and in a total volume of approximately 2 mL. Twenty one and 42 days after the initial injection, booster doses, which are identical to priming doses, except that Freund's incomplete adjuvant is used, are administered in the same way. Fifteen days after the last injection, animal serum is recovered, decomplexed, and filtered through a 0.45 μ m membrane.

2.G.2. Mouse hyperimmune ascites fluid

Ten mice are injected subcutaneously with 10-50 μ g of a purified fusion polypeptide as obtained in 2.E.1. or 2.E.2., in the presence of Freund's complete adjuvant and in a volume of approximately 200 μ L. Seven and 14 days after the
5 initial injection, booster doses, which are identical to the priming doses, except that Freund's incomplete adjuvant is used, are administered in the same way. Twenty one and 28 days after the initial infection, mice receive 50 μ g of the antigen alone intraperitoneally. On day 21, mice are also injected intraperitoneally with sarcoma 180/TG cells CM26684 (Lennette *et al.*, *Diagnostic Procedures for*
10 *Viral, Rickettsial, and Chlamydial Infections*, 5th Ed. Washington DC, American Public Health Association, 1979). Ascites fluid is collected 10-13 days after the last injection.

EXAMPLE 3: Methods for producing transcriptional fusions lacking His-tags

15 Methods for amplification and cloning of DNA encoding the polypeptides of the invention as transcriptional fusions lacking His-tags are described as follows. Two PCR primers for each clone are designed based upon the sequences of the polynucleotides that encode them (SEQ ID NOs:1-97 (odd numbers), 99, and 100). These primers can be used to amplify DNA encoding the polypeptides
20 of the invention from any *Helicobacter pylori* strain, including, for example, ORV2001 and the strain deposited as ATCC deposit number 43579, as well as from other *Helicobacter* species.

The N-terminal primers are designed to include the ribosome binding site of the target gene, the ATG start site, and any leader sequence and cleavage site. The

N-terminal primers can include a 5' clamp and a restriction endonuclease recognition site, such as that for *Bam*HI (GGATCC), which facilitates subsequent cloning. Similarly, the C-terminal primers can include a restriction endonuclease recognition site, such as that for *Xho*I (CTCGAG), which can be used in subsequent cloning, and a TAA stop codon.

Amplification of genes encoding the polypeptides of the invention is carried out using Thermalase DNA Polymerase under the conditions described above in Example 2. Alternatively, Vent DNA polymerase (New England Biolabs), Pwo DNA polymerase (Boehringer Mannheim), or Taq DNA polymerase (Appligene) can be used, according to instructions provided by the manufacturers.

A single PCR product for each clone is amplified and cloned into appropriately cleaved pET 24 (*e.g.*, *Bam*HI-*Xho*I cleaved pET 24), resulting in construction of a transcriptional fusion that permits expression of the proteins without His-tags. The expressed products can be purified as denatured proteins that are refolded by dialysis into 1 M arginine.

Cloning into pET 24 allows transcription of the genes from the T7 promoter, which is supplied by the vector, but relies upon binding of the RNA-specific DNA polymerase to the intrinsic ribosome binding sites of the genes, and thereby expression of the complete ORF. The amplification, digestion, and cloning protocols are as described above for constructing translational fusions.

EXAMPLE 4: Purification of the polypeptides of the invention by immunoaffinity

4.A. Purification of specific IgGs

An immune serum, as prepared in section 2.G., is applied to a protein A
5 Sepharose Fast Flow column (Pharmacia) equilibrated in 100 mM Tris-HCl (pH 8.0). The resin is washed by applying 10 column volumes of 100 mM Tris-HCl and 10 volumes of 10 mM Tris-HCl (pH 8.0) to the column. IgG antibodies are eluted with 0.1 M glycine buffer (pH 3.0) and are collected as 5 mL fractions to which is added 0.25 mL 1 M Tris-HCl (pH 8.0). The optical density of the eluate
10 is measured at 280 nm and the fractions containing the IgG antibodies are pooled, dialyzed against 50 mM Tris-HCl (pH 8.0), and, if necessary, stored frozen at -70°C.

4.B. Preparation of the column

An appropriate amount of CNBr-activated Sepharose 4B gel (1 g of dried
15 gel provides for approximately 3.5 mL of hydrated gel; gel capacity is from 5 to 10 mg coupled IgG/mL of gel) manufactured by Pharmacia (17-0430-01) is suspended in 1 mM HCl buffer and washed with a buchner by adding small quantities of 1 mM HCl buffer. The total volume of buffer is 200 mL per gram of gel.

20 Purified IgG antibodies are dialyzed for 4 hours at $20\pm 5^{\circ}\text{C}$ against 50 volumes of 500 mM sodium phosphate buffer (pH 7.5). The antibodies are then diluted in 500 mM phosphate buffer (pH 7.5) to a final concentration of 3 mg/mL.

IgG antibodies are mixed with the gel overnight at $5\pm 3^{\circ}\text{C}$. The gel is packed into a chromatography column and is washed with 2 column volumes of 500 mM phosphate buffer (pH 7.5), and 1 column volume of 50 mM sodium phosphate buffer, containing 500 mM NaCl (pH 7.5). The gel is then transferred to a tube, mixed with 100 mM ethanolamine (pH 7.5) for 4 hours at room temperature, and washed twice with 2 column volumes of PBS. The gel is then stored in 1/10,000 PBS/merthiolate. The amount of IgG antibodies coupled to the gel is determined by measuring the optical density (OD) at 280 nm of the IgG solution and the direct eluate, plus washings.

4.C. Adsorption and elution of the antigen

An antigen solution in 50 mM Tris-HCl (pH 8.0), 2 mM EDTA, for example, the supernatant obtained in 3.E. or the solubilized pellet obtained in 3.E., after centrifugation and filtration through a $0.45\ \mu\text{m}$ membrane, is applied to a column equilibrated with 50 mM Tris-HCl (pH 8.0), 2 mM EDTA, at a flow rate of about 10 mL/hour. The column is then washed with 20 volumes of 50 mM Tris-HCl (pH 8.0), 2 mM EDTA. Alternatively, adsorption can be achieved by mixing overnight at $5\pm 3^{\circ}\text{C}$.

The adsorbed gel is washed with 2 to 6 volumes of 10 mM sodium phosphate buffer (pH 6.8) and the antigen is eluted with 100 mM glycine buffer (pH 2.5). The eluate is recovered in 3 mL fractions, to each of which is added 150 μL of 1 M sodium phosphate buffer (pH 8.0). Absorption is measured at 280 nm for each fraction; those fractions containing the antigen are pooled and stored at -20°C .

Other embodiments are within the following claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: Haas, Rainer et al.

5 (ii) TITLE OF THE INVENTION: Helicobacter Polypeptides
and Corresponding Polynucleotides

(iii) NUMBER OF SEQUENCES: 112

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Clark & Elbing LLP

(B) STREET: 585 Commercial Street

10 (C) CITY: Boston

(D) STATE: MA

(E) COUNTRY: USA

(F) ZIP: 02109

(v) COMPUTER READABLE FORM:

15 (A) MEDIUM TYPE: Diskette

(B) COMPUTER: IBM Compatible

(C) OPERATING SYSTEM: DOS

(D) SOFTWARE: FastSEQ for Windows Version 2.0

(vi) CURRENT APPLICATION DATA:

20 (A) APPLICATION NUMBER:

(B) FILING DATE: 01-APR-1997

(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

25 (A) NAME: Clark, Paul T.

(B) REGISTRATION NUMBER: 30,175

(C) REFERENCE/DOCKET NUMBER: 06132/040001

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 617-428-0200

(B) TELEFAX: 617-428-7045

30 (C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 989 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence
(B) LOCATION: 71...940
(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

	CTATGACGAT TGTCTCGCTT TTAGAAAACA CTCTAATCGC TTTTGAAAAA CAACAAAGGA	60
10	AGGGATTTTA ATG AAA TTT TTA CGC TCT GTT TAT GCA TTT TGC TCC AGT	109
	Met Lys Phe Leu Arg Ser Val Tyr Ala Phe Cys Ser Ser	
	1 5 10	
	TGG GTA GGG ACG ATT GTT ATT GTG CTG TTG GTT ATC TTT TTT ATC GCG	157
	Trp Val Gly Thr Ile Val Ile Val Leu Leu Val Ile Phe Phe Ile Ala	
15	15 20 25	
	CAA GCC TTT ATC ATT CCC TCT CGC TCT ATG GTT GGC ACG CTC TAT GAG	205
	Gln Ala Phe Ile Ile Pro Ser Arg Ser Met Val Gly Thr Leu Tyr Glu	
	30 35 40 45	
	GGC GAC ATG CTC TTT GTC AAA AAG TTT TCT TAC GGC ATA CCC ATT CCT	253
20	Gly Asp Met Leu Phe Val Lys Lys Phe Ser Tyr Gly Ile Pro Ile Pro	
	50 55 60	
	AAA ATC CCA TGG ATT GAG CTT CCT GTT ATG CCT GAT TTT AAA AAT AAC	301
	Lys Ile Pro Trp Ile Glu Leu Pro Val Met Pro Asp Phe Lys Asn Asn	
	65 70 75	
25	GGA CAT TTG ATA GAG GGG GAT CGC CCT AAG CGT GGC GAA GTG GTG GTG	349
	Gly His Leu Ile Glu Gly Asp Arg Pro Lys Arg Gly Glu Val Val Val	
	80 85 90	
	TTT ATC CCT CCC CAT GAA AAA AAG TCT TAC TAT GTT AAA AGG AAT TTT	397
	Phe Ile Pro Pro His Glu Lys Lys Ser Tyr Tyr Val Lys Arg Asn Phe	
30	95 100 105	
	GCC ATT GGA GGC GAT GAG GTG TTG TTC ACT AAT GAG GGT TTT TAT TTG	445
	Ala Ile Gly Gly Asp Glu Val Leu Phe Thr Asn Glu Gly Phe Tyr Leu	
	110 115 120 125	
	CAC CCT TTT GAG AGC GAC ACG GAC AAA AAT TAC ATC GCT AAA CAT TAC	493
35	His Pro Phe Glu Ser Asp Thr Asp Lys Asn Tyr Ile Ala Lys His Tyr	
	130 135 140	
	CCT AAC GCC ATG ACA AAA GAA TTT ATG GGT AAA ATT TTT GTT TTA AAC	541

	Pro	Asn	Ala	Met	Thr	Lys	Glu	Phe	Met	Gly	Lys	Ile	Phe	Val	Leu	Asn	
				145					150					155			
	CCT	TAT	AAA	AAT	GAG	CAT	CCG	GGT	ATC	CAT	TAC	CAA	AAA	GAC	AAT	GAA	589
5	Pro	Tyr	Lys	Asn	Glu	His	Pro	Gly	Ile	His	Tyr	Gln	Lys	Asp	Asn	Glu	
			160					165					170				
	ACC	TTC	CAC	TTA	ATG	GAG	CAA	TTA	GCC	ACT	CAA	GGC	GCA	GAA	GCT	AAT	637
	Thr	Phe	His	Leu	Met	Glu	Gln	Leu	Ala	Thr	Gln	Gly	Ala	Glu	Ala	Asn	
			175					180					185				
	ATC	AGC	ATG	CAA	CTC	ATT	CAA	ATG	GAG	GGC	GAA	AAG	GTG	TTT	TAT	AAG	685
10	Ile	Ser	Met	Gln	Leu	Ile	Gln	Met	Glu	Gly	Glu	Lys	Val	Phe	Tyr	Lys	
			190				195				200					205	
	AAA	ATC	AAT	GAC	GAT	GAA	TTT	TTC	ATG	ATC	GGC	GAC	AAC	AGA	GAC	AAT	733
	Lys	Ile	Asn	Asp	Asp	Glu	Phe	Phe	Met	Ile	Gly	Asp	Asn	Arg	Asp	Asn	
				210						215					220		
15	TCT	AGC	GAC	TCG	CGC	TTT	TGG	GGG	AGT	GTG	GCT	TAT	AAA	AAC	ATC	GTG	781
	Ser	Ser	Asp	Ser	Arg	Phe	Trp	Gly	Ser	Val	Ala	Tyr	Lys	Asn	Ile	Val	
				225				230						235			
	GGT	TCG	CCA	TGG	TTT	GTT	TAT	TTC	AGT	TTG	AGT	TTA	AAA	AAT	AGC	CTA	829
20	Gly	Ser	Pro	Trp	Phe	Val	Tyr	Phe	Ser	Leu	Ser	Leu	Lys	Asn	Ser	Leu	
			240					245					250				
	GAA	ATG	GAT	GCA	GAA	AAT	AAC	CCT	AAA	AAA	CGC	TAT	CTG	GTG	CGT	TGG	877
	Glu	Met	Asp	Ala	Glu	Asn	Asn	Pro	Lys	Lys	Arg	Tyr	Leu	Val	Arg	Trp	
			255				260					265					
	GAA	CGC	ATG	TTT	AAA	AGC	GTT	GGA	GGC	TTA	GAA	AAA	ATC	ATT	AAA	AAA	925
25	Glu	Arg	Met	Phe	Lys	Ser	Val	Gly	Gly	Leu	Glu	Lys	Ile	Ile	Lys	Lys	
			270				275				280				285		
	GAA	AAC	GCA	ACG	CAT	TAAGGTTTTT	TGTGCAATTT	TTTGATTCT	CTTTAGAAAG	T							981
	Glu	Asn	Ala	Thr	His												
				290													
30	TTTATTAC																989

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 290 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Lys Phe Leu Arg Ser Val Tyr Ala Phe Cys Ser Ser Trp Val Gly
1 5 10 15
5 Thr Ile Val Ile Val Leu Leu Val Ile Phe Phe Ile Ala Gln Ala Phe
20 25 30
Ile Ile Pro Ser Arg Ser Met Val Gly Thr Leu Tyr Glu Gly Asp Met
35 40 45
10 Leu Phe Val Lys Lys Phe Ser Tyr Gly Ile Pro Ile Pro Lys Ile Pro
50 55 60
Trp Ile Glu Leu Pro Val Met Pro Asp Phe Lys Asn Asn Gly His Leu
65 70 75 80
Ile Glu Gly Asp Arg Pro Lys Arg Gly Glu Val Val Val Phe Ile Pro
85 90 95
15 Pro His Glu Lys Lys Ser Tyr Tyr Val Lys Arg Asn Phe Ala Ile Gly
100 105 110
Gly Asp Glu Val Leu Phe Thr Asn Glu Gly Phe Tyr Leu His Pro Phe
115 120 125
20 Glu Ser Asp Thr Asp Lys Asn Tyr Ile Ala Lys His Tyr Pro Asn Ala
130 135 140
Met Thr Lys Glu Phe Met Gly Lys Ile Phe Val Leu Asn Pro Tyr Lys
145 150 155 160
Asn Glu His Pro Gly Ile His Tyr Gln Lys Asp Asn Glu Thr Phe His
165 170 175
25 Leu Met Glu Gln Leu Ala Thr Gln Gly Ala Glu Ala Asn Ile Ser Met
180 185 190
Gln Leu Ile Gln Met Glu Gly Glu Lys Val Phe Tyr Lys Lys Ile Asn
195 200 205
Asp Asp Glu Phe Phe Met Ile Gly Asp Asn Arg Asp Asn Ser Ser Asp
210 215 220
30 Ser Arg Phe Trp Gly Ser Val Ala Tyr Lys Asn Ile Val Gly Ser Pro
225 230 235 240
Trp Phe Val Tyr Phe Ser Leu Ser Leu Lys Asn Ser Leu Glu Met Asp
245 250 255
35 Ala Glu Asn Asn Pro Lys Lys Arg Tyr Leu Val Arg Trp Glu Arg Met
260 265 270
Phe Lys Ser Val Gly Gly Leu Glu Lys Ile Ile Lys Lys Glu Asn Ala
275 280 285
Thr His
40 290

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- 45 (A) LENGTH: 514 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 112...471
- (D) OTHER INFORMATION:

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

	GGATTTT	TTA	GAGCTCT	TTAG	TCAATGATAA	TGTGGTAGAA	ACGATTGAAA	AAGGCTTT	TGT	60							
	GATAGGTTT	GGAGCGGGG	ATATTACCTA	TCAATTAAGA	GGCGAAATGT	A	ATG	GGT		117							
							Met	Gly									
10									1								
	GCA	GTG	GTT	GTT	TTA	TTT	TTA	ACG	CTG	GTT	TTA	TTG	TTT	TTA	GTT	TTA	165
	Ala	Val	Val	Val	Leu	Phe	Leu	Thr	Leu	Val	Leu	Leu	Phe	Leu	Val	Leu	
			5						10							15	
	AGG	GAT	TTT	GGT	TTA	GCA	AGC	CCC	AAA	CAA	AAG	ATT	TTA	GCT	TTT	TTA	213
15	Arg	Asp	Phe	Gly	Leu	Ala	Ser	Pro	Lys	Gln	Lys	Ile	Leu	Ala	Phe	Leu	
		20					25				30						
	ATC	GTA	GGG	ATT	ATA	GGA	GCG	AGC	ATC	AGC	GTT	TAT	ACT	TAC	AAG	CAA	261
	Ile	Val	Gly	Ile	Ile	Gly	Ala	Ser	Ile	Ser	Val	Tyr	Thr	Tyr	Lys	Gln	
		35				40					45					50	
20	AAC	CAA	CAA	AAC	CAA	CAA	GAG	ATC	GCT	TTG	CAA	AGA	GCG	TTT	TTA	AGG	309
	Asn	Gln	Gln	Asn	Gln	Gln	Glu	Ile	Ala	Leu	Gln	Arg	Ala	Phe	Leu	Arg	
				55						60					65		
	GGG	GAA	ACC	TTG	TTG	TGT	AAA	GGC	ATT	AAA	GTC	AAT	AAC	CAA	ACC	TTT	357
25	Gly	Glu	Thr	Leu	Leu	Cys	Lys	Gly	Ile	Lys	Val	Asn	Asn	Gln	Thr	Phe	
			70						75					80			
	AAT	TTA	GTG	AGC	GGA	ACT	TTA	AGC	TTT	TTA	GGC	AAA	AAA	CAA	ACC	CCT	405
	Asn	Leu	Val	Ser	Gly	Thr	Leu	Ser	Phe	Leu	Gly	Lys	Lys	Gln	Thr	Pro	
			85					90					95				
	ATG	AAA	GAC	GTT	CTT	GTG	GAT	TTG	GAT	TCT	TGT	CAG	ACG	CTC	CAA	AAA	453
30	Met	Lys	Asp	Val	Leu	Val	Asp	Leu	Asp	Ser	Cys	Gln	Thr	Leu	Gln	Lys	
		100					105				110						
	GAT	CCC	TTA	ATC	CAA	CCC	TAATGATGAA	TAATAATAAT	ACCCCACCCA	AACCCCTA							509
	Asp	Pro	Leu	Ile	Gln	Pro											
		115				120											
35	GAAGA																514

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 120 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
10  Met Gly Ala Val Val Val Leu Phe Leu Thr Leu Val Leu Leu Phe Leu
    1          5          10          15
    Val Leu Arg Asp Phe Gly Leu Ala Ser Pro Lys Gln Lys Ile Leu Ala
        20          25          30
    Phe Leu Ile Val Gly Ile Ile Gly Ala Ser Ile Ser Val Tyr Thr Tyr
        35          40          45
15  Lys Gln Asn Gln Gln Asn Gln Gln Glu Ile Ala Leu Gln Arg Ala Phe
    50          55          60
    Leu Arg Gly Glu Thr Leu Leu Cys Lys Gly Ile Lys Val Asn Asn Gln
    65          70          75          80
    Thr Phe Asn Leu Val Ser Gly Thr Leu Ser Phe Leu Gly Lys Lys Gln
    85          90          95
20  Thr Pro Met Lys Asp Val Leu Val Asp Leu Asp Ser Cys Gln Thr Leu
    100          105          110
    Gln Lys Asp Pro Leu Ile Gln Pro
        115          120
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1233 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
(B) LOCATION: 135...1049
(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```
40  GTTTTAAATT TAATATTCAT TAAGCTTTTG TGGCTATTCC ATTTTAATTT TGTTTTTCAT      60
    TAAACCCAA TCTAAATCT TATTTTATG ATAAATACC TAATCATAAT ATCAAATCTT      120
    AAACCAACGA AACC ATG AAA AAA GCT CTC TTA CTA ACT CTC TCT CTC TCG      170
    Met Lys Lys Ala Leu Leu Leu Thr Leu Ser Leu Ser
        1          5          10
```

	TTC TGG CTC CAC GCT GAA AGG AAT GGA TTT TAT TTA GGT TTA AAT TTT	218
	Phe Trp Leu His Ala Glu Arg Asn Gly Phe Tyr Leu Gly Leu Asn Phe	
	15 20 25	
	CTA GAA GGA AGC TAT ATT AAA GGA CAA GGT AGC ATC GGC AAA AAA GCT	266
5	Leu Glu Gly Ser Tyr Ile Lys Gly Gln Gly Ser Ile Gly Lys Lys Ala	
	30 35 40	
	TCA GCA GAA AAC GCC TTA AAT GAA GCG ATC AAT AAC GCA AAA AAT TCA	314
	Ser Ala Glu Asn Ala Leu Asn Glu Ala Ile Asn Asn Ala Lys Asn Ser	
	45 50 55 60	
10	TTA TTC CCT AAC ACA AAA GCC ATA AGA GAT GCA CAA AAC GCC TTA AAT	362
	Leu Phe Pro Asn Thr Lys Ala Ile Arg Asp Ala Gln Asn Ala Leu Asn	
	65 70 75	
	GCA GTG AAA GAT TCA AAC AAA ATC GCT AGC CGA TTC GCA GGA AAT GGT	410
	Ala Val Lys Asp Ser Asn Lys Ile Ala Ser Arg Phe Ala Gly Asn Gly	
15	80 85 90	
	GGA TCG GGC GGT CTT TTT AAT GAG CTC AGC TTT GGG TAT AAA TAT TTT	458
	Gly Ser Gly Gly Leu Phe Asn Glu Leu Ser Phe Gly Tyr Lys Tyr Phe	
	95 100 105	
	TTG GGT AAA AAA AGG ATT ATA GGG TTT AGG CAC TCT CTT TTT TTC GGT	506
20	Leu Gly Lys Lys Arg Ile Ile Gly Phe Arg His Ser Leu Phe Phe Gly	
	110 115 120	
	TAC CAA CTT GGT GGC GTT GGT TCT GTT CCT GGT AGC GGT TTA ATC GTT	554
	Tyr Gln Leu Gly Gly Val Gly Ser Val Pro Gly Ser Gly Leu Ile Val	
	125 130 135 140	
25	TTT TTA CCC TAT GGT TTC AAT ACG GAT TTG CTC ATT AAT TGG ACT AAC	602
	Phe Leu Pro Tyr Gly Phe Asn Thr Asp Leu Leu Ile Asn Trp Thr Asn	
	145 150 155	
	GAT AAG CGA GCG TCC CAA AAA TAT GTT GAA CGA AGG GTA AAA GGG CTC	650
	Asp Lys Arg Ala Ser Gln Lys Tyr Val Glu Arg Arg Val Lys Gly Leu	
30	160 165 170	
	TCT ATA TTT TAC AAA GAT ATG ACC GGC AGA ACG CTA GAC GCT AAT ACA	698
	Ser Ile Phe Tyr Lys Asp Met Thr Gly Arg Thr Leu Asp Ala Asn Thr	
	175 180 185	
	TTA AAA AAA GCA TCA AGG CAT GTA TTT AGA AAA TCT TCA GGG CTT GTG	746
35	Leu Lys Lys Ala Ser Arg His Val Phe Arg Lys Ser Ser Gly Leu Val	
	190 195 200	
	ATT GGC ATG GAA CTA GGG GGT AGC ACT TGG TTT GCA AGT AAC AAT CTC	794
	Ile Gly Met Glu Leu Gly Gly Ser Thr Trp Phe Ala Ser Asn Asn Leu	

205	210	215	220	
ACC CCT TTC AAT CAA GTC AAG AGT CGC ACG ATT TTT CAG TTG CAA GGA	842			
Thr Pro Phe Asn Gln Val Lys Ser Arg Thr Ile Phe Gln Leu Gln Gly				
225 230 235				

5	AAA TTT GGC GTT CGT TGG AAT AAT GAT GAA TAC GAT ATT GAT CGC TAT	890
	Lys Phe Gly Val Arg Trp Asn Asn Asp Glu Tyr Asp Ile Asp Arg Tyr	
	240 245 250	

10	GGC GAT GAA ATC TAT CTT GGA GGT TCT AGT GTT GAA TTA GGG GTT AAA	938
	Gly Asp Glu Ile Tyr Leu Gly Gly Ser Ser Val Glu Leu Gly Val Lys	
	255 260 265	

	GTG CCA GCG TTT AAA GTC AAT TAC TAT AGC GAT GAT TAT GGG GAT AAA	986
	Val Pro Ala Phe Lys Val Asn Tyr Tyr Ser Asp Asp Tyr Gly Asp Lys	
	270 275 280	

15	TTG GAT TAT AAA AGA GTG GTG AGC GTT TAT CTT AAC TAT ACA TAT AAC	1034
	Leu Asp Tyr Lys Arg Val Val Ser Val Tyr Leu Asn Tyr Thr Tyr Asn	
	285 290 295 300	

	TTT AAA AAC AAA CAT TAAAACACGC TTTTACCGC TCTTTAGTTG GTTTTTTAAA A	1090
	Phe Lys Asn Lys His	
	305	

20	AACCTTATTT TTTATTAGCT TGAAACTCTT CAAAGCCTTT TTTTCTCAAT TGGCATGCCG	1150
	GGCATTATATC GCAACCATAA CCATAAGCAT GCAAAATCTT TCGCTCTCCT TGATAGCAGG	1210
	TGTGCGTTTC TTTGATGACT AAA	1233

(2) INFORMATION FOR SEQ ID NO:6:

25	(i) SEQUENCE CHARACTERISTICS:
	(A) LENGTH: 305 amino acids
	(B) TYPE: amino acid
	(C) STRANDEDNESS: single
	(D) TOPOLOGY: linear

30	(ii) MOLECULE TYPE: protein
	(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

	Met Lys Lys Ala Leu Leu Thr Leu Ser Leu Ser Phe Trp Leu His
	1 5 10 15
35	Ala Glu Arg Asn Gly Phe Tyr Leu Gly Leu Asn Phe Leu Glu Gly Ser
	20 25 30
	Tyr Ile Lys Gly Gln Gly Ser Ile Gly Lys Lys Ala Ser Ala Glu Asn
	35 40 45
	Ala Leu Asn Glu Ala Ile Asn Asn Ala Lys Asn Ser Leu Phe Pro Asn

	AATGACGGCT	CTAAACCAAA	CGATTTGACT	TCTCCAAAAG	AAGCCTCTCA	AGAATCTCAA	60
	AAAAATGAAG	CTCCAAAAAA	TGAAGTTCAA	AGAAATGAAG	CTCAAAAAGA	AACCCCCCAA	120
	TCCAATCAAA	CGCCTAAAGA	A ATG AAA	GTC AAG	TCC ATT	TCT TAT GTC GGG	171
5			Met Lys	Val Lys	Ser Ile	Ser Tyr Val Gly	10
			1		5		
	CTT TCT TAC	ATG TCT	GAC ATG	CTC GCT	AAT GAA	ATT GTA AAG	ATT CGT
	Leu Ser Tyr	Met Ser	Asp Met	Leu Ala	Asn Glu	Ile Val Lys	Ile Arg
		15		20		25	219
	GTG GGC GAT	ATT GTG	GAT TCT	AAA AAA	ATA GAC	ACC GCT GTT	TTG GCT
10	Val Gly Asp	Ile Val	Asp Ser	Lys Lys	Ile Asp	Thr Ala Val	Leu Ala
		30		35		40	267
	TTG TTC AAT	CAA GGG	TAT TTT	AAA GAC	GTT TAT	GCC ACT TTT	GAA GGC
	Leu Phe Asn	Gln Gly	Tyr Phe	Lys Asp	Val Tyr	Ala Thr Phe	Glu Gly
		45		50		55	315
15	GGC ATA TTA	GAG TTT	CAT TTT	GAT GAA	AAA GCC	AGG ATT GCC	GGG GTA
	Gly Ile Leu	Glu Phe	His Phe	Asp Glu	Lys Ala	Arg Ile Ala	Gly Val
		60		65		70	363
	GAA ATC AAG	GGT TAT	GGG ACT	GAA AAG	GAA AAA	GAC GGC TTA	AAA TCC
	Glu Ile Lys	Gly Tyr	Gly Thr	Glu Lys	Glu Lys	Asp Gly Leu	Lys Ser
20	75		80		85		90
	CAA ATG GGG	ATC AAA	AAG GGC	GAC ACC	TTT GAT	GAG CAA AAA	TTA GAG
	Gln Met Gly	Ile Lys	Lys Gly	Asp Thr	Phe Asp	Glu Gln Lys	Leu Glu
		95		100		105	459
	CAT GCT AAA	ACG GCT	TTA AAA	ACC GCT	TTA GAG	GGG CAG GGC	TAT TAT
25	His Ala Lys	Thr Ala	Leu Lys	Thr Ala	Leu Glu	Gly Gln Gly	Tyr Tyr
		110		115		120	507
	GGG AGC GTG	GTG GAG	GTG CGC	ACA GAA	AAG GTC	AGT GAG GGT	GCA TTA
	Gly Ser Val	Val Glu	Val Arg	Thr Glu	Lys Val	Ser Glu Gly	Ala Leu
		125		130		135	555
30	TTG ATC GTG	TTT GAT	GTG AAT	AGG GGG	GAT AGC	ATT TAT ATC	AAA CAA
	Leu Ile Val	Phe Asp	Val Asn	Arg Gly	Asp Ser	Ile Tyr Ile	Lys Gln
		140		145		150	603
	TCC ATT TAT	GAG GGA	AGC GCG	AAA TTA	AAA CGC	CGC ATG ATT	GAA TCT
	Ser Ile Tyr	Glu Gly	Ser Ala	Lys Leu	Lys Arg	Arg Met Ile	Glu Ser
35	155		160		165		170
	TTG AGT GCG	AAC AAG	CAA CGA	GAT TTC	ATG GGC	TGG ATG TGG	GGC TTG
	Leu Ser Ala	Asn Lys	Gln Arg	Asp Phe	Met Gly	Trp Met Trp	Gly Leu
		175		180		185	699

	AAT GAC GGG AAA TTG CGT TTA GAT CAA CTA GAA TAC GAT TCT ATG CGT	747
	Asn Asp Gly Lys Leu Arg Leu Asp Gln Leu Glu Tyr Asp Ser Met Arg	
	190 195 200	
5	ATC CAA GAT GTG TAT ATG CGT AGG GGT TAC TTA GAC GCT CAT ATT TCT	795
	Ile Gln Asp Val Tyr Met Arg Arg Gly Tyr Leu Asp Ala His Ile Ser	
	205 210 215	
	TCG CCT TTT TTG AAA ACG GAT TTT TCT ACC CAT GAC GCT AAG CTT CAT	843
	Ser Pro Phe Leu Lys Thr Asp Phe Ser Thr His Asp Ala Lys Leu His	
	220 225 230	
10	TAT AAA GTC AAA GAG GGG ATC CAA TAC AGG ATT TCA GAC ATT TTA ATA	891
	Tyr Lys Val Lys Glu Gly Ile Gln Tyr Arg Ile Ser Asp Ile Leu Ile	
	235 240 245 250	
15	GAG ATT GAC AAC CCG GTA GTC CCC TTA AAA ACC TTA GAA AAA GCG CTT	939
	Glu Ile Asp Asn Pro Val Val Pro Leu Lys Thr Leu Glu Lys Ala Leu	
	255 260 265	
	AAA GTG AAA AGG AAA GAT GTC TTT AAT ATT GAG CAT TTA AGA GCG GAT	987
	Lys Val Lys Arg Lys Asp Val Phe Asn Ile Glu His Leu Arg Ala Asp	
	270 275 280	
20	GCG CAA ATT TTA AAA ACC GAA ATC GCC GAT AAG GGT TAT GCG TTT GCG	1035
	Ala Gln Ile Leu Lys Thr Glu Ile Ala Asp Lys Gly Tyr Ala Phe Ala	
	285 290 295	
	GTG GTG AAG CCA GAC TTG GAT AAA GAT GAA AAA AAC GGG CTT GTG AAA	1083
	Val Val Lys Pro Asp Leu Asp Lys Asp Glu Lys Asn Gly Leu Val Lys	
	300 305 310	
25	GTC ATT TAT CGT ATT GAA GTG GGC GAT ATG GTG TAT ATC AAT GAT GTC	1131
	Val Ile Tyr Arg Ile Glu Val Gly Asp Met Val Tyr Ile Asn Asp Val	
	315 320 325 330	
30	ATC ATT TCA GGG AAC CAG CGC ACG AGC GAT AGG ATC ATT AGA AGG GAG	1179
	Ile Ile Ser Gly Asn Gln Arg Thr Ser Asp Arg Ile Ile Arg Arg Glu	
	335 340 345	
	TTA TTG TTA GGG CCT AAG GAT AAA TAC AAC TTG ACC AAA CTG AGA AAT	1227
	Leu Leu Leu Gly Pro Lys Asp Lys Tyr Asn Leu Thr Lys Leu Arg Asn	
	350 355 360	
35	TCC GAA AAT TCT TTA AGG CGT TTA GGA TTC TTC TCT AAA GTC AAA ATT	1275
	Ser Glu Asn Ser Leu Arg Arg Leu Gly Phe Phe Ser Lys Val Lys Ile	
	365 370 375	
	GAA GAA AAA AGG GTT AAT AGC TCA CTC ATG GAT TTA TTA GTG AGC GTA	1323
	Glu Glu Lys Arg Val Asn Ser Ser Leu Met Asp Leu Leu Val Ser Val	

	380	385	390	
	GAA GAG GGG CGT ACT GGG CAG TTG CAA TTT GGG TTA GGC TAT GGC TCT			1371
	Glu Glu Gly Arg Thr Gly Gln Leu Gln Phe Gly Leu Gly Tyr Gly Ser			
	395	400	405	410
5	TAT GGA GGG CTT ATG CTT AAT GGG AGC GTG AGC GAA AGA AAC CTT TTT			1419
	Tyr Gly Gly Leu Met Leu Asn Gly Ser Val Ser Glu Arg Asn Leu Phe			
		415	420	425
	GGC ACA GGG CAA AGC ATG AGC TTG TAT GCT AAC ATC GCT ACA GGG GGG			1467
10	Gly Thr Gly Gln Ser Met Ser Leu Tyr Ala Asn Ile Ala Thr Gly Gly			
		430	435	440
	GGT AGA TCT TAT CCG GGC ATG CCA AAA GGA GCG GGG CGT ATG TTT GCC			1515
	Gly Arg Ser Tyr Pro Gly Met Pro Lys Gly Ala Gly Arg Met Phe Ala			
		445	450	455
	GGG AAT TTG AGC TTG ACT AAT CCA AGG ATT TTT GAC AGC TGG TAT AGC			1563
15	Gly Asn Leu Ser Leu Thr Asn Pro Arg Ile Phe Asp Ser Trp Tyr Ser			
		460	465	470
	TCT ACG ATC AAC CTT TAT GCG GAT TAC AGG ATA AGC TAC CAA TAC ATC			1611
	Ser Thr Ile Asn Leu Tyr Ala Asp Tyr Arg Ile Ser Tyr Gln Tyr Ile			
		475	480	485
20	CAA CAA GGC GGG GGC TTT GGG GTG AAT GTC GGG CGC ATG CTG GGT AAT			1659
	Gln Gln Gly Gly Gly Phe Gly Val Asn Val Gly Arg Met Leu Gly Asn			
		495	500	505
	AGA ACC CAT GTG AGC TTA GGG TAT AAC TTG AAT GTT ACC AAA CTC CTT			1707
25	Arg Thr His Val Ser Leu Gly Tyr Asn Leu Asn Val Thr Lys Leu Leu			
		510	515	520
	GGT TTC AGC AGC CCT TTA TAC AAC CGC TAC TAT TCC TCT GTT AAT GAA			1755
	Gly Phe Ser Ser Pro Leu Tyr Asn Arg Tyr Tyr Ser Ser Val Asn Glu			
		525	530	535
	GTG GTT TCT CCA AGG CAA TGT TCT ACC CCC GCA TCG GTG ATT ATC AAT			1803
30	Val Val Ser Pro Arg Gln Cys Ser Thr Pro Ala Ser Val Ile Ile Asn			
		540	545	550
	CGC TTA TCA GGC GGT AAA ACC CCC TTA CAA CCT GAA AGC TGT TCT AGT			1851
	Arg Leu Ser Gly Gly Lys Thr Pro Leu Gln Pro Glu Ser Cys Ser Ser			
		555	560	565
35	CCT GGA GCG ATC ACC ACT TCA CCA GAA ATA AGA GGT ATT TGG GAT AGG			1899
	Pro Gly Ala Ile Thr Thr Ser Pro Glu Ile Arg Gly Ile Trp Asp Arg			
		575	580	585

	GAT TAC CAT ACG CCT ATC ACC AGC TCT TTC ACC CTT GAT GTG AGC TAT	1947
	Asp Tyr His Thr Pro Ile Thr Ser Ser Phe Thr Leu Asp Val Ser Tyr	
	590 595 600	
	GAC AAC ACC GAT GAT TAT TAC TTC CCT AGA AAT GGG GTT ATC TTT AGT	1995
5	Asp Asn Thr Asp Asp Tyr Tyr Phe Pro Arg Asn Gly Val Ile Phe Ser	
	605 610 615	
	TCC TAT GCG ACG ATG TCT GGC TTG CCA AGC TCT GGC ACG CTC AAT TCT	2043
	Ser Tyr Ala Thr Met Ser Gly Leu Pro Ser Ser Gly Thr Leu Asn Ser	
	620 625 630	
10	TGG AAC GGG TTA GGC GGG AAT GTC CGT AAC ACC AAA GTT TAT GGT AAA	2091
	Trp Asn Gly Leu Gly Gly Asn Val Arg Asn Thr Lys Val Tyr Gly Lys	
	635 640 645 650	
	TTC GCC GCT TAC CAC CAT TTG CAA AAA TAT TTA TTG ATA GAT TTG ATC	2139
	Phe Ala Ala Tyr His His Leu Gln Lys Tyr Leu Leu Ile Asp Leu Ile	
15	655 660 665	
	GCT CGC TTT AAA ACG CAA GGA GGT TAT ATC TTT AGG TAT AAC ACC GAT	2187
	Ala Arg Phe Lys Thr Gln Gly Gly Tyr Ile Phe Arg Tyr Asn Thr Asp	
	670 675 680	
	GAT TAC TTG CCC TTA AAC TCC ACC TTC TAC ATG GGG GGC GTA ACC ACG	2235
20	Asp Tyr Leu Pro Leu Asn Ser Thr Phe Tyr Met Gly Gly Val Thr Thr	
	685 690 695	
	GTG AGA GGC TTT AGG AAC GGA TCG GTT ACT CCT AAA GAT GAG TTT GGC	2283
	Val Arg Gly Phe Arg Asn Gly Ser Val Thr Pro Lys Asp Glu Phe Gly	
	700 705 710	
25	TTG TGG CTT GGA GGC GAT GGG ATT TTT ACC GCT TCT ACT GAA TTG AGC	2331
	Leu Trp Leu Gly Gly Asp Gly Ile Phe Thr Ala Ser Thr Glu Leu Ser	
	715 720 725 730	
	TAT GGG GTG CTA AAG GCG GCT AAA ATG CGC TTA GCG TGG TTT TTT GAC	2379
	Tyr Gly Val Leu Lys Ala Ala Lys Met Arg Leu Ala Trp Phe Phe Asp	
30	735 740 745	
	TTT GGT TTC TTA ACC TTT AAA ACC CCA ACT AGA GGG AGT TTT TTC TAT	2427
	Phe Gly Phe Leu Thr Phe Lys Thr Pro Thr Arg Gly Ser Phe Phe Tyr	
	750 755 760	
	AAC GCT CCT GTT ACG ACA GCG AAT TTT AAA GAT TAT GGC GTT ATA GGG	2475
35	Asn Ala Pro Val Thr Thr Ala Asn Phe Lys Asp Tyr Gly Val Ile Gly	
	765 770 775	
	GCT GGG TTT GAA AGA GCG ACT TGG AGG GCT TCC ACA GGC TTG CAG ATT	2523
	Ala Gly Phe Glu Arg Ala Thr Trp Arg Ala Ser Thr Gly Leu Gln Ile	

780

785

790

GAA TGG ATT TCG CCC ATG GGG CCT TTG GTG TTG ATT TTC CCT ATA GCG 2571
 Glu Trp Ile Ser Pro Met Gly Pro Leu Val Leu Ile Phe Pro Ile Ala
 795 800 805 810

5 TTT TTC AAC CAA TGG GGC GAT GGC AAT GGC AAG AAA TGT AAA GGG CTA 2619
 Phe Phe Asn Gln Trp Gly Asp Gly Asn Gly Lys Lys Cys Lys Gly Leu
 815 820 825

TGC TTC AAC CCT AAC ATG GAC GAT TAC ACG CAA CAC TTT GAA TTT TCT 2667
 Cys Phe Asn Pro Asn Met Asp Asp Tyr Thr Gln His Phe Glu Phe Ser
 10 830 835 840

ATG GGA ACA AGG TTT TAAAATGCGC ATCAACAGAG AAGAAATTTT GGATTTAATG A 2723
 Met Gly Thr Arg Phe
 845

AAAACGCGCC CTTGAAAGAA TTGGGGCAAA GGGCTTTGAG GGTGAAGCAA CGCTTGCACC 2783
 15 CTGAAACTTT GACGACTTTT ATTGTGGATA GGAATATCAA TTACACCAAT ATTTGTTTTG 2843
 TGGATTGCAA GTTTTGCGCG TTCAAACGCA CCTTAAAGA AAAAGACGCC TATGTGTTGA 2903
 GCTATGAAGA AATTGATCAA AAGATTGAAG AATTGCTCGC TATTGGCGGC ACGCAGATCC 2963
 TTTTTC AAGG GGGGGTGCAC CCGCAGCTAA AGATTGATTA TTATGAGAA 3012

(2) INFORMATION FOR SEQ ID NO:8:

20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 847 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

25 (ii) MOLECULE TYPE: protein
 (v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Lys Val Lys Ser Ile Ser Tyr Val Gly Leu Ser Tyr Met Ser Asp
 1 5 10 15
 30 Met Leu Ala Asn Glu Ile Val Lys Ile Arg Val Gly Asp Ile Val Asp
 20 25 30
 Ser Lys Lys Ile Asp Thr Ala Val Leu Ala Leu Phe Asn Gln Gly Tyr
 35 40 45
 Phe Lys Asp Val Tyr Ala Thr Phe Glu Gly Gly Ile Leu Glu Phe His
 35 50 55 60
 Phe Asp Glu Lys Ala Arg Ile Ala Gly Val Glu Ile Lys Gly Tyr Gly
 65 70 75 80
 Thr Glu Lys Glu Lys Asp Gly Leu Lys Ser Gln Met Gly Ile Lys Lys
 85 90 95
 40 Gly Asp Thr Phe Asp Glu Gln Lys Leu Glu His Ala Lys Thr Ala Leu

			100				105				110					
	Lys	Thr	Ala	Leu	Glu	Gly	Gln	Gly	Tyr	Tyr	Gly	Ser	Val	Val	Glu	Val
			115					120					125			
5	Arg	Thr	Glu	Lys	Val	Ser	Glu	Gly	Ala	Leu	Leu	Ile	Val	Phe	Asp	Val
			130					135					140			
	Asn	Arg	Gly	Asp	Ser	Ile	Tyr	Ile	Lys	Gln	Ser	Ile	Tyr	Glu	Gly	Ser
	145					150					155				160	
	Ala	Lys	Leu	Lys	Arg	Arg	Met	Ile	Glu	Ser	Leu	Ser	Ala	Asn	Lys	Gln
				165						170					175	
10	Arg	Asp	Phe	Met	Gly	Trp	Met	Trp	Gly	Leu	Asn	Asp	Gly	Lys	Leu	Arg
			180						185					190		
	Leu	Asp	Gln	Leu	Glu	Tyr	Asp	Ser	Met	Arg	Ile	Gln	Asp	Val	Tyr	Met
			195					200					205			
15	Arg	Arg	Gly	Tyr	Leu	Asp	Ala	His	Ile	Ser	Ser	Pro	Phe	Leu	Lys	Thr
			210					215					220			
	Asp	Phe	Ser	Thr	His	Asp	Ala	Lys	Leu	His	Tyr	Lys	Val	Lys	Glu	Gly
	225					230					235				240	
	Ile	Gln	Tyr	Arg	Ile	Ser	Asp	Ile	Leu	Ile	Glu	Ile	Asp	Asn	Pro	Val
				245						250					255	
20	Val	Pro	Leu	Lys	Thr	Leu	Glu	Lys	Ala	Leu	Lys	Val	Lys	Arg	Lys	Asp
			260						265					270		
	Val	Phe	Asn	Ile	Glu	His	Leu	Arg	Ala	Asp	Ala	Gln	Ile	Leu	Lys	Thr
			275					280					285			
25	Glu	Ile	Ala	Asp	Lys	Gly	Tyr	Ala	Phe	Ala	Val	Val	Lys	Pro	Asp	Leu
			290				295					300				
	Asp	Lys	Asp	Glu	Lys	Asn	Gly	Leu	Val	Lys	Val	Ile	Tyr	Arg	Ile	Glu
	305					310					315				320	
	Val	Gly	Asp	Met	Val	Tyr	Ile	Asn	Asp	Val	Ile	Ile	Ser	Gly	Asn	Gln
				325						330					335	
30	Arg	Thr	Ser	Asp	Arg	Ile	Ile	Arg	Arg	Glu	Leu	Leu	Leu	Gly	Pro	Lys
			340						345					350		
	Asp	Lys	Tyr	Asn	Leu	Thr	Lys	Leu	Arg	Asn	Ser	Glu	Asn	Ser	Leu	Arg
			355					360					365			
35	Arg	Leu	Gly	Phe	Phe	Ser	Lys	Val	Lys	Ile	Glu	Glu	Lys	Arg	Val	Asn
			370				375					380				
	Ser	Ser	Leu	Met	Asp	Leu	Leu	Val	Ser	Val	Glu	Glu	Gly	Arg	Thr	Gly
	385				390					395					400	
	Gln	Leu	Gln	Phe	Gly	Leu	Gly	Tyr	Gly	Ser	Tyr	Gly	Gly	Leu	Met	Leu
				405					410						415	
40	Asn	Gly	Ser	Val	Ser	Glu	Arg	Asn	Leu	Phe	Gly	Thr	Gly	Gln	Ser	Met
			420						425					430		
	Ser	Leu	Tyr	Ala	Asn	Ile	Ala	Thr	Gly	Gly	Gly	Arg	Ser	Tyr	Pro	Gly
			435					440					445			
45	Met	Pro	Lys	Gly	Ala	Gly	Arg	Met	Phe	Ala	Gly	Asn	Leu	Ser	Leu	Thr
			450				455					460				
	Asn	Pro	Arg	Ile	Phe	Asp	Ser	Trp	Tyr	Ser	Ser	Thr	Ile	Asn	Leu	Tyr
	465				470					475					480	
	Ala	Asp	Tyr	Arg	Ile	Ser	Tyr	Gln	Tyr	Ile	Gln	Gln	Gly	Gly	Gly	Phe
				485					490						495	
50	Gly	Val	Asn	Val	Gly	Arg	Met	Leu	Gly	Asn	Arg	Thr	His	Val	Ser	Leu

				500					505					510			
	Gly	Tyr	Asn	Leu	Asn	Val	Thr	Lys	Leu	Leu	Gly	Phe	Ser	Ser	Pro	Leu	
			515					520					525				
5	Tyr	Asn	Arg	Tyr	Tyr	Ser	Ser	Val	Asn	Glu	Val	Val	Ser	Pro	Arg	Gln	
		530					535					540					
	Cys	Ser	Thr	Pro	Ala	Ser	Val	Ile	Ile	Asn	Arg	Leu	Ser	Gly	Gly	Lys	
	545					550				555						560	
	Thr	Pro	Leu	Gln	Pro	Glu	Ser	Cys	Ser	Ser	Pro	Gly	Ala	Ile	Thr	Thr	
					565					570						575	
10	Ser	Pro	Glu	Ile	Arg	Gly	Ile	Trp	Asp	Arg	Asp	Tyr	His	Thr	Pro	Ile	
				580					585					590			
	Thr	Ser	Ser	Phe	Thr	Leu	Asp	Val	Ser	Tyr	Asp	Asn	Thr	Asp	Asp	Tyr	
		595					600					605					
	Tyr	Phe	Pro	Arg	Asn	Gly	Val	Ile	Phe	Ser	Ser	Tyr	Ala	Thr	Met	Ser	
15		610					615					620					
	Gly	Leu	Pro	Ser	Ser	Gly	Thr	Leu	Asn	Ser	Trp	Asn	Gly	Leu	Gly	Gly	
	625					630					635					640	
	Asn	Val	Arg	Asn	Thr	Lys	Val	Tyr	Gly	Lys	Phe	Ala	Ala	Tyr	His	His	
					645					650						655	
20	Leu	Gln	Lys	Tyr	Leu	Leu	Ile	Asp	Leu	Ile	Ala	Arg	Phe	Lys	Thr	Gln	
				660					665					670			
	Gly	Gly	Tyr	Ile	Phe	Arg	Tyr	Asn	Thr	Asp	Asp	Tyr	Leu	Pro	Leu	Asn	
		675					680					685					
	Ser	Thr	Phe	Tyr	Met	Gly	Gly	Val	Thr	Thr	Val	Arg	Gly	Phe	Arg	Asn	
25		690					695					700					
	Gly	Ser	Val	Thr	Pro	Lys	Asp	Glu	Phe	Gly	Leu	Trp	Leu	Gly	Gly	Asp	
	705					710					715					720	
	Gly	Ile	Phe	Thr	Ala	Ser	Thr	Glu	Leu	Ser	Tyr	Gly	Val	Leu	Lys	Ala	
					725					730						735	
30	Ala	Lys	Met	Arg	Leu	Ala	Trp	Phe	Phe	Asp	Phe	Gly	Phe	Leu	Thr	Phe	
			740					745						750			
	Lys	Thr	Pro	Thr	Arg	Gly	Ser	Phe	Phe	Tyr	Asn	Ala	Pro	Val	Thr	Thr	
		755					760						765				
	Ala	Asn	Phe	Lys	Asp	Tyr	Gly	Val	Ile	Gly	Ala	Gly	Phe	Glu	Arg	Ala	
35		770					775					780					
	Thr	Trp	Arg	Ala	Ser	Thr	Gly	Leu	Gln	Ile	Glu	Trp	Ile	Ser	Pro	Met	
	785					790					795					800	
	Gly	Pro	Leu	Val	Leu	Ile	Phe	Pro	Ile	Ala	Phe	Phe	Asn	Gln	Trp	Gly	
					805					810						815	
40	Asp	Gly	Asn	Gly	Lys	Lys	Cys	Lys	Gly	Leu	Cys	Phe	Asn	Pro	Asn	Met	
			820						825					830			
	Asp	Asp	Tyr	Thr	Gln	His	Phe	Glu	Phe	Ser	Met	Gly	Thr	Arg	Phe		
		835					840						845				

(2) INFORMATION FOR SEQ ID NO:9:

- 45 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1032 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence

(B) LOCATION: 149...913

5 (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

	ATGTTTTGTG	TTGCAAAAAC	AAAACAGACC	AATAAAGGCA	TCACTTTTAA	AAGCGTTGTT	60
	TAGGGGGGTT	TGTTATTGG	TGTTTGATTA	GAATAGGGTT	GTTTTTAATT	TTCTTTTAAG	120
10	AGGAGTTTTT	ACTTTTTTAA	GGGTTTTT	ATG GAT ATT	TAT GCG TTA	TAT ATA	172
				Met Asp Ile	Tyr Ala Leu	Tyr Ile	
				1		5	
	GCG ATA GGG	CTT TTT ACT	GGC ATT CTA	TCA GGG ATT	TTT GGC ATT	GGT	220
	Ala Ile Gly	Leu Phe Thr	Gly Ile Leu	Ser Gly Ile	Phe Gly Ile	Gly	
15	10		15		20		
	GGG GGG TTG	ATC ATT GTC	CCT ATC ATG	CTC GCA ACC	GGG CAT TCT	TTT	268
	Gly Gly Leu	Ile Ile Val	Pro Ile Met	Leu Ala Thr	Gly His Ser	Phe	
	25		30		35	40	
	GAA GAA TCC	ATT GGG ATT	TCC ATT TTG	CAA ATG GCG	CTT TCA TCG	TTC	316
20	Glu Glu Ser	Ile Gly Ile	Ser Ile Leu	Gln Met Ala	Leu Ser Ser	Phe	
		45		50		55	
	GTG GGC TCT	GTT TTG AAT	TTC AAA AAA	AAA TCG CTT	GAT TTT TCT	TTA	364
	Val Gly Ser	Val Leu Asn	Phe Lys Lys	Lys Ser Leu	Asp Phe Ser	Leu	
		60		65		70	
25	GGC TTG TTG	ATA GGG GCA	GGG GGG CTG	ATA GGG GCG	AGT TTT AGC	GGA	412
	Gly Leu Leu	Ile Gly Ala	Gly Gly Leu	Ile Gly Ala	Ser Phe Ser	Gly	
		75		80		85	
	TTT GTT TTA	AAA ATC GTT	TCC AGT AAA	ATT TTA ATG	GTT ATT TTC	GCG	460
	Phe Val Leu	Lys Ile Val	Ser Ser Lys	Ile Leu Met	Val Ile Phe	Ala	
30	90		95		100		
	CTT TTA GTC	GTG TAT TCT	ATG ATC CAA	TTT GTT TTG	AAA CCC AAA	AAA	508
	Leu Leu Val	Val Tyr Ser	Met Ile Gln	Phe Val Leu	Lys Pro Lys	Lys	
	105		110		115	120	
	AAA GAT TTG	ATA GCG GAT	ACT AAA CGC	TAT CAT CTG	CAA GGT TTG	AAA	556
35	Lys Asp Leu	Ile Ala Asp	Thr Lys Arg	Tyr His Leu	Gln Gly Leu	Lys	
		125		130		135	
	TTA TTT TTA	ATT GGC ACG	CTC ACA GGG	TTT TTT GCT	ATC ACT TTA	GGG	604

Leu Phe Leu Ile Gly Thr Leu Thr Gly Phe Phe Ala Ile Thr Leu Gly
140 145 150

ATT GGT GGG GGG ATG CTC ATG GTG CCT TTG ATG CAT TAT TTT TTA GGG 652
Ile Gly Gly Gly Met Leu Met Val Pro Leu Met His Tyr Phe Leu Gly
5 155 160 165

TAT GAT TCT AAA AAA TGC GTG GCT CTA GGG TTA TTT TTC ATC TTG TTT 700
Tyr Asp Ser Lys Lys Cys Val Ala Leu Gly Leu Phe Phe Ile Leu Phe
170 175 180

TCT TCT ATT TCA GGA GCT TTT TCT TTA ATG TAT CAC CAC ATC ATC AAT 748
10 Ser Ser Ile Ser Gly Ala Phe Ser Leu Met Tyr His His Ile Ile Asn
185 190 195 200

AAA GAA GTG CTC TTA GCA GGG GCG ATT GTG GGA TTA GGA TCT GTT ATG 796
Lys Glu Val Leu Leu Ala Gly Ala Ile Val Gly Leu Gly Ser Val Met
205 210 215

GGC GTG AGC ATT GGG ATT AAA TGG ATC ATG GGG CTT TTG AAT GAA AAA 844
15 Gly Val Ser Ile Gly Ile Lys Trp Ile Met Gly Leu Leu Asn Glu Lys
220 225 230

ATG CAT AAA GCT TTG ATT TTA GGG GTG TAT GGT TTG TCG CTA TTG ATT 892
20 Met His Lys Ala Leu Ile Leu Gly Val Tyr Gly Leu Ser Leu Leu Ile
235 240 245

GTT TTA TAC AAA CTC TTT TTT TAATTGATGG TTTTATACCA CTACTATTTT AAGA 947
Val Leu Tyr Lys Leu Phe Phe
250 255

CCCCTAAGAG TTTCCCTTTA GAGTATTTGC ATTTGTGCGC TAATGAGAGC CATTATTGA 1007
25 GATTGGATTT TGATGCGGCC AATTT 1032

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- 30 (A) LENGTH: 255 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

35 Met Asp Ile Tyr Ala Leu Tyr Ile Ala Ile Gly Leu Phe Thr Gly Ile
1 5 10 15
Leu Ser Gly Ile Phe Gly Ile Gly Gly Gly Leu Ile Ile Val Pro Ile

	1	5	10	15	
	TTA GCC AGC TTG GGG GTT TTG TGG GGC AAT AGC GCT GAA AAA ACG CCT				158
	Leu Ala Ser Leu Gly Val Leu Trp Gly Asn Ser Ala Glu Lys Thr Pro	20	25	30	
5	TTT TTT AAA ACG AAA AAC CAC ATT TAT CTA GGT TTT AGG CTA GGC ACA				206
	Phe Phe Lys Thr Lys Asn His Ile Tyr Leu Gly Phe Arg Leu Gly Thr	35	40	45	
	GGA GCC AAT GTG CAC ACG AGC ATG TGG CAA CAA GCC TAT AAA GAC AAC				254
10	Gly Ala Asn Val His Thr Ser Met Trp Gln Gln Ala Tyr Lys Asp Asn	50	55	60	
	CCC ACC TGC CCT GGT AGC GTG TGT TAT GGC GAG AAA TTA GAA GCC CAT				302
	Pro Thr Cys Pro Gly Ser Val Cys Tyr Gly Glu Lys Leu Glu Ala His	65	70	75	
	TAT CAG GGG GGT AAA AAC CTG TCT TAT ACC GGG CAA ATA GGC GAT GAA				350
15	Tyr Gln Gly Gly Lys Asn Leu Ser Tyr Thr Gly Gln Ile Gly Asp Glu	80	85	90	95
	ATA GCT TTT GAT AAA CAC CAT ATT TTA GGC TTA AGG GTG TGG GGG GAT				398
	Ile Ala Phe Asp Lys His His Ile Leu Gly Leu Arg Val Trp Gly Asp	100	105	110	
20	GTA GAA TAC GCT AAA GCG CAA TTA GGT CAA AAA GTG GGG GGT AAT ACC				446
	Val Glu Tyr Ala Lys Ala Gln Leu Gly Gln Lys Val Gly Gly Asn Thr	115	120	125	
	CTT TTA TCC CAA GCC AAT TAT GAC CCA AAC GCG ATT AAA ACC TAC GAT				494
25	Leu Leu Ser Gln Ala Asn Tyr Asp Pro Asn Ala Ile Lys Thr Tyr Asp	130	135	140	
	TCT GCT TCA AAC ACT CAA GGC CCT TTA GTT TTG CAA AAA ACC CCA AGC				542
	Ser Ala Ser Asn Thr Gln Gly Pro Leu Val Leu Gln Lys Thr Pro Ser	145	150	155	
	CCT CAA AAC TTC CTT TTC AAT AAC GGG CAT TTC ATG GCG TTT GGT TTG				590
30	Pro Gln Asn Phe Leu Phe Asn Asn Gly His Phe Met Ala Phe Gly Leu	160	165	170	175
	AAC GTG AAT GTG TTT GTT AAC CTC CCT ATA GAC ACC CTT TTA AAA CTC				638
	Asn Val Asn Val Phe Val Asn Leu Pro Ile Asp Thr Leu Leu Lys Leu	180	185	190	
35	GCT TTA AAA ACA GAA AAA ATG CTG TTT TTT AAA ATA GGC GTG TTT GGT				686
	Ala Leu Lys Thr Glu Lys Met Leu Phe Phe Lys Ile Gly Val Phe Gly	195	200	205	

GGG GGC GGG GTG GAA TAC GCA ATA TTA TGG AGT CCT AAC TAT CAA AAT 734
 Gly Gly Gly Val Glu Tyr Ala Ile Leu Trp Ser Pro Asn Tyr Gln Asn
 210 215 220

CAA AAC ACG AAA CAA GGC GAT AAA TTT TTT GCA GCG GGT GGG GGG TTT 782
 5 Gln Asn Thr Lys Gln Gly Asp Lys Phe Phe Ala Ala Gly Gly Gly Phe
 225 230 235

TTT GTG AAT TTT GGG GGT TCT TTG TAT ATA GGC AAA CGC AAC CGC TTC 830
 Phe Val Asn Phe Gly Gly Ser Leu Tyr Ile Gly Lys Arg Asn Arg Phe
 240 245 250 255

10 AAT GTG GGG TTA AAA ATC CCT TAC TAT AGC TTG AGC GCG CAA AGT TGG 878
 Asn Val Gly Leu Lys Ile Pro Tyr Tyr Ser Leu Ser Ala Gln Ser Trp
 260 265 270

AAA AAC TTT GGC TCT AGC AAT GTG TGG CAG CAA CAA ACG ATC CGA CAA 926
 Lys Asn Phe Gly Ser Ser Asn Val Trp Gln Gln Gln Thr Ile Arg Gln
 15 275 280 285

AAC TTC AGC GTT TTT AGG AAT AAA GAA GTT TTT GTC AGC TAC GCG TTC 974
 Asn Phe Ser Val Phe Arg Asn Lys Glu Val Phe Val Ser Tyr Ala Phe
 290 295 300

TTG TTT TAGTTTGGAT TCGTTCTCAT TAAACACTGA TGATAAAATT CAAAAGATGG TT 1032
 20 Leu Phe
 305

TTATCGTTAC AAAATTCAAC ATTTTC 1057

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- 25 (A) LENGTH: 305 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

- 30 (v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Met Gly Arg Ile Glu Ser Lys Lys Arg Leu Lys Ala Leu Val Phe Leu
 1 5 10 15
 Ala Ser Leu Gly Val Leu Trp Gly Asn Ser Ala Glu Lys Thr Pro Phe
 20 25 30
 Phe Lys Thr Lys Asn His Ile Tyr Leu Gly Phe Arg Leu Gly Thr Gly
 35 40 45
 Ala Asn Val His Thr Ser Met Trp Gln Gln Ala Tyr Lys Asp Asn Pro

	TATTAGTTGG	TTTAATACGC	TATAATCTGT	GTGCCAACAT	TGTGTGGCTC	AAATCATTTT	60
	TAAAAGGGGT	TTTATA	ATG	GAA	AAC	AAC	112
			Met	Glu	Asn	Asn	
			1		5	10	
5	GTT	TTG	CGC	AAG	TTT	TTA	160
	Val	Leu	Arg	Lys	Phe	Leu	
		15			20		
						25	
	GGA	GGA	TTG	AAT	ATG	GAA	208
	Gly	Gly	Leu	Asn	Met	Glu	
10		30			35		
						40	
	CCA	ATC	ATG	AAT	ATT	TTG	256
	Pro	Ile	Met	Asn	Ile	Leu	
		45			50		
						55	
						60	
	TCG	CTC	ATT	AAC	GCT	CTA	304
15	Ser	Leu	Ile	Asn	Ala	Leu	
				65			
						70	
						75	
	GGA	AAA	CCC	ATC	ACT	CAG	352
	Gly	Lys	Pro	Ile	Thr	Gln	
				80			
						85	
						90	
20	GGC	TTG	ATT	TTA	TGG	GAC	400
	Gly	Leu	Ile	Leu	Trp	Asp	
		95				100	
						105	
	AAT	ACC	TTG	GAA	AGC	ATT	448
	Asn	Thr	Leu	Glu	Ser	Ile	
25		110				115	
							120
	CTT	GAT	GAA	AAA	GAG	GCT	496
	Leu	Asp	Glu	Lys	Glu	Ala	
		125				130	
						135	
						140	
	ACT	TCT	GGT	AGG	GTT	CAA	547
30	Thr	Ser	Gly	Arg	Val	Gln	
						145	
							150
	AAAAATGGAA	TATCCCAACG	ATTTTCGTTT	TCACCAACAC	ACAAGAAAAA	GCCGGCGATG	607
	CCTTTGTTAA	AAAAACT					624

(2) INFORMATION FOR SEQ ID NO:14:

- 35 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 153 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

```
5 Met Glu Asn Asn Glu Asn His Glu Lys Leu Asn Gly Val Leu Arg Lys
  1          5          10          15
Phe Leu Gly Asp Ala Phe Thr Leu Asp Gly Lys Glu Gly Gly Leu Asn
  20          25          30
10 Met Glu Lys Leu Arg Glu Ala Ile Lys Lys Glu Lys Pro Ile Met Asn
  35          40          45
Ile Leu Leu Met Gly Ala Thr Gly Val Gly Lys Ser Ser Leu Ile Asn
  50          55          60
Ala Leu Phe Gly Lys Glu Val Ala Lys Ala Gly Val Gly Lys Pro Ile
  65          70          75          80
15 Thr Gln His Leu Glu Lys Tyr Val Asp Glu Glu Lys Gly Leu Ile Leu
  85          90          95
Trp Asp Thr Lys Gly Ile Glu Asp Lys Asp Tyr Glu Asn Thr Leu Glu
  100          105          110
Ser Ile Lys Lys Glu Met Glu Asp Ser Phe Lys Thr Leu Asp Glu Lys
  115          120          125
20 Glu Ala Ile Asp Val Ala Tyr Leu Cys Val Lys Glu Thr Ser Gly Arg
  130          135          140
Val Gln Glu Arg Glu Arg Glu Ser Tyr
  145          150
```

25 (2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1083 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

30 (D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence

(B) LOCATION: 155...1033

(D) OTHER INFORMATION:

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

```
40 GATGTTGTTA AGTCGTTGTT TATTATGTTA CACTAAAAGC TTAAATAAAA GGGCATAAGG      60
GATAAAGGGA GTGTTAGTAG ATAGTTTAA TAGGGTTATT GACTATATTA GGGTTTCTGT      120
AACCAAACAG TGCAATTTCA GGTGTCAGTA TTGC ATG CCT GCT ACG CCA TTA AAT      175
                               Met Pro Ala Thr Pro Leu Asn
                               1          5
```

	TTT TTT GAT AAT GAA GAA TTA TTG CCT TTG GAT AAT GTT TTA GAA TTT	223
	Phe Phe Asp Asn Glu Glu Leu Leu Pro Leu Asp Asn Val Leu Glu Phe	
	10 15 20	
	CTC AAA ATC GCC ATT GAT GAG GGC GTT AAA AAA ATT AGA ATC ACG GGT	271
5	Leu Lys Ile Ala Ile Asp Glu Gly Val Lys Lys Ile Arg Ile Thr Gly	
	25 30 35	
	GGG GAG CCG CTA TTA CGC AAA GGC TTA GAT GAA TTT ATC GCT AAA TTG	319
	Gly Glu Pro Leu Leu Arg Lys Gly Leu Asp Glu Phe Ile Ala Lys Leu	
	40 45 50 55	
10	CAC GCT TAC AAT AAA GAA GTG GAG TTA GTT TTA AGC ACT AAT GGT TTT	367
	His Ala Tyr Asn Lys Glu Val Glu Leu Val Leu Ser Thr Asn Gly Phe	
	60 65 70	
	TTA CTC AAA AAA ATG GCT AAG GAT TTA AAA AAT GCC GGG TTA GCG CAA	415
15	Leu Leu Lys Lys Met Ala Lys Asp Leu Lys Asn Ala Gly Leu Ala Gln	
	75 80 85	
	GTG AAT GTT TCA TTG GAT TCT TTA AAA AGC GAT AGG GTT TTA AAA ATC	463
	Val Asn Val Ser Leu Asp Ser Leu Lys Ser Asp Arg Val Leu Lys Ile	
	90 95 100	
	TCT CAA AAA GAC GCT CTT AAA AAC ACG CTA GAA GGG ATT GAA GAG TCT	511
20	Ser Gln Lys Asp Ala Leu Lys Asn Thr Leu Glu Gly Ile Glu Glu Ser	
	105 110 115	
	TTG AAA GTG GGT TTA AAA CTC AAA TTA AAC ACG GTT GTG ATA AAA AGC	559
	Leu Lys Val Gly Leu Lys Leu Lys Leu Asn Thr Val Val Ile Lys Ser	
	120 125 130 135	
25	GTT AAT GAT GAT GAA ATC TTA GAG CTT TTA GAA TAC GCA AAA AAT AGG	607
	Val Asn Asp Asp Glu Ile Leu Glu Leu Leu Glu Tyr Ala Lys Asn Arg	
	140 145 150	
	CAT ATA CAA ATC CGC TAC ATT GAA TTT ATG GAA AAC ACG CAT GCT AAA	655
30	His Ile Gln Ile Arg Tyr Ile Glu Phe Met Glu Asn Thr His Ala Lys	
	155 160 165	
	AGT TTG GTT AAA GGC TTG AAA GAG CGA GAA ATT TTA GAT TTG ATC GCT	703
	Ser Leu Val Lys Gly Leu Lys Glu Arg Glu Ile Leu Asp Leu Ile Ala	
	170 175 180	
	CAA AAA TAT CAA ATC ATT GAG GCA GAA AAA CCC AAA CAA GGG TCT TCT	751
35	Gln Lys Tyr Gln Ile Ile Glu Ala Glu Lys Pro Lys Gln Gly Ser Ser	
	185 190 195	
	AAA ATC TAC ACG CTA GAA AAT GGC TAT CAA TTT GGC ATT ATC GCT CCG	799
	Lys Ile Tyr Thr Leu Glu Asn Gly Tyr Gln Phe Gly Ile Ile Ala Pro	

	200	205	210	215	
	CAT AGC GAT GAT TTT TGC CAA TCT TGC AAT CGT ATC CGT TTG GCT TCT				847
	His Ser Asp Asp Phe Cys Gln Ser Cys Asn Arg Ile Arg Leu Ala Ser				
		220	225	230	
5	GAT GGT AAG ATT TGC CCA TGT TTA TAC TAT CAA GAC GCC ATA GAC GCT				895
	Asp Gly Lys Ile Cys Pro Cys Leu Tyr Tyr Gln Asp Ala Ile Asp Ala				
		235	240	245	
	AAA GAG GCG ATC ATC AAT AAG GAT ACA AAA AAT ATA AAA AGG CTT TTA				943
	Lys Glu Ala Ile Ile Asn Lys Asp Thr Lys Asn Ile Lys Arg Leu Leu				
10		250	255	260	
	AAG CAA TCT GTC ATC AAT AAA CCA GAA AAA AAC ATG TGG AAT GAT AAA				991
	Lys Gln Ser Val Ile Asn Lys Pro Glu Lys Asn Met Trp Asn Asp Lys				
		265	270	275	
	AAC AGC GAA ACT CCC ACA AGG GCG TTT TAC TAC ACA GGG GGG TAGGGGAGT				1042
15	Asn Ser Glu Thr Pro Thr Arg Ala Phe Tyr Tyr Thr Gly Gly				
		280	285	290	
	AAAATATTTA TTATTTTAAA CCTTTTATT AAAAATAAGG C				1083

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 293 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

	Met Pro Ala Thr Pro Leu Asn Phe Phe Asp Asn Glu Glu Leu Leu Pro	
	1 5 10 15	
	Leu Asp Asn Val Leu Glu Phe Leu Lys Ile Ala Ile Asp Glu Gly Val	
30		
	Lys Lys Ile Arg Ile Thr Gly Gly Glu Pro Leu Leu Arg Lys Gly Leu	
	35 40 45	
	Asp Glu Phe Ile Ala Lys Leu His Ala Tyr Asn Lys Glu Val Glu Leu	
	50 55 60	
35	Val Leu Ser Thr Asn Gly Phe Leu Leu Lys Lys Met Ala Lys Asp Leu	
	65 70 75 80	
	Lys Asn Ala Gly Leu Ala Gln Val Asn Val Ser Leu Asp Ser Leu Lys	
	85 90 95	
	Ser Asp Arg Val Leu Lys Ile Ser Gln Lys Asp Ala Leu Lys Asn Thr	

	Thr	Glu	Thr	Gln	Arg	Leu	Arg	Phe	Trp	Trp	Trp	Ser	Ser	Phe	Gly	Arg	
				20					25					30			
	GGT	TTA	GCG	TTC	AAC	TCT	TGG	ATT	TAT	GAG	TTT	TTT	GCG	ACG	GTG	CCT	264
5	Gly	Leu	Ala	Phe	Asn	Ser	Trp	Ile	Tyr	Glu	Phe	Phe	Ala	Thr	Val	Pro	
			35					40					45				
	TAT	TTG	AAA	AAG	GGA	GGC	AAT	CCT	AAT	AAC	AGC	AAC	GAT	TTC	ATC	AAT	312
	Tyr	Leu	Lys	Lys	Gly	Gly	Asn	Pro	Asn	Asn	Ser	Asn	Asp	Phe	Ile	Asn	
			50				55					60					
	TAT	GGC	TGG	CAT	GGA	ATC	ACC	ACA	ACC	TAT	TCT	TAT	AAA	GGT	TTA	GAC	360
10	Tyr	Gly	Trp	His	Gly	Ile	Thr	Thr	Thr	Tyr	Ser	Tyr	Lys	Gly	Leu	Asp	
	65					70					75				80		
	GCT	CAA	TTT	TTT	TAT	TAT	TTT	GCG	CCT	AAG	ACT	TAT	AAC	GCT	CCT	GGC	408
	Ala	Gln	Phe	Phe	Tyr	Tyr	Phe	Ala	Pro	Lys	Thr	Tyr	Asn	Ala	Pro	Gly	
					85					90				95			
15	TTT	AAG	CTG	GTC	TAT	GAC	ACG	AAT	AGG	AAT	TTT	CAA	AAT	GTA	GGC	TTT	456
	Phe	Lys	Leu	Val	Tyr	Asp	Thr	Asn	Arg	Asn	Phe	Gln	Asn	Val	Gly	Phe	
				100					105					110			
	CGC	TCT	CAA	AGC	ATG	ATC	ATG	ACA	ACC	TTT	CCT	TTA	TAC	TAT	AGA	GGG	504
20	Arg	Ser	Gln	Ser	Met	Ile	Met	Thr	Thr	Phe	Pro	Leu	Tyr	Tyr	Arg	Gly	
			115					120					125				
	TGG	TAT	AAC	CCA	GAG	ACA	AAC	ACT	TAT	AGT	TTA	GAA	GAC	AGC	ACG	CCT	552
	Trp	Tyr	Asn	Pro	Glu	Thr	Asn	Thr	Tyr	Ser	Leu	Glu	Asp	Ser	Thr	Pro	
			130				135					140					
25	CAT	GGC	TCG	TTG	TTG	GGG	AGG	AAT	GGC	GTT	ACT	TTA	AAT	ATC	CGC	CAG	600
	His	Gly	Ser	Leu	Leu	Gly	Arg	Asn	Gly	Val	Thr	Leu	Asn	Ile	Arg	Gln	
						150				155					160		
	GTT	TTT	TGG	TGG	GAT	AAT	TTC	AAC	TGG	TCC	ATT	GGC	TTT	TAT	AAC	ACC	648
	Val	Phe	Trp	Trp	Asp	Asn	Phe	Asn	Trp	Ser	Ile	Gly	Phe	Tyr	Asn	Thr	
					165					170				175			
30	TTT	GGC	AAT	TCG	GAC	GCT	TTT	TTA	GGC	TCT	CAC	ACG	ATG	CCA	AGG	GGT	696
	Phe	Gly	Asn	Ser	Asp	Ala	Phe	Leu	Gly	Ser	His	Thr	Met	Pro	Arg	Gly	
				180					185					190			
	AAT	AAC	ACT	TCC	TAT	ATC	GGT	AGT	GAA	ATC	TCC	ATA	ACG	ACT	AGG	CAT	744
35	Asn	Asn	Thr	Ser	Tyr	Ile	Gly	Ser	Glu	Ile	Ser	Ile	Thr	Thr	Arg	His	
			195					200					205				
	GCC	GGA	ATG	ATT	GGC	TAT	GAT	TTT	TGG	GAT	AAT	ACG	GCT	TAT	GAT	GGG	792
	Ala	Gly	Met	Ile	Gly	Tyr	Asp	Phe	Trp	Asp	Asn	Thr	Ala	Tyr	Asp	Gly	
		210					215					220					

Gly Leu Ala Phe Asn Ser Trp Ile Tyr Glu Phe Phe Ala Thr Val Pro
 35 40 45
 Tyr Leu Lys Lys Gly Gly Asn Pro Asn Asn Ser Asn Asp Phe Ile Asn
 50 55 60
 5 Tyr Gly Trp His Gly Ile Thr Thr Thr Tyr Ser Tyr Lys Gly Leu Asp
 65 70 75 80
 Ala Gln Phe Phe Tyr Thr Phe Ala Pro Lys Thr Tyr Asn Ala Pro Gly
 85 90 95
 Phe Lys Leu Val Tyr Asp Thr Asn Arg Asn Phe Gln Asn Val Gly Phe
 10 100 105 110
 Arg Ser Gln Ser Met Ile Met Thr Thr Phe Pro Leu Tyr Tyr Arg Gly
 115 120 125
 Trp Tyr Asn Pro Glu Thr Asn Thr Tyr Ser Leu Glu Asp Ser Thr Pro
 130 135 140
 15 His Gly Ser Leu Leu Gly Arg Asn Gly Val Thr Leu Asn Ile Arg Gln
 145 150 155 160
 Val Phe Trp Trp Asp Asn Phe Asn Trp Ser Ile Gly Phe Tyr Asn Thr
 165 170 175
 Phe Gly Asn Ser Asp Ala Phe Leu Gly Ser His Thr Met Pro Arg Gly
 180 185 190
 20 Asn Asn Thr Ser Tyr Ile Gly Ser Glu Ile Ser Ile Thr Thr Arg His
 195 200 205
 Ala Gly Met Ile Gly Tyr Asp Phe Trp Asp Asn Thr Ala Tyr Asp Gly
 210 215 220
 25 Leu Ala Asp Ala Ile Thr Asn Ala Asn Thr Phe Thr Phe Tyr Thr Ser
 225 230 235 240
 Val Gly Gly Ile His Lys Arg Phe Ala Trp His Val Phe Gly Arg Val
 245 250 255
 Ser His Ala Asn Lys Asn Ala Leu Gly Gln Val Gly Arg Ala Asn Glu
 260 265 270
 30 Tyr Ser Leu Gln Phe Asn Ala Ser Tyr Ala Phe Thr Glu Ser Ile Leu
 275 280 285
 Leu Asn Phe Arg Ile Thr Tyr Tyr Gly Ala Arg Ile Asn Lys Gly Tyr
 290 295 300
 35 Gln Ala Gly Tyr Phe Gly Ala Pro Lys Phe Asn Asn Pro Asp Gly Asp
 305 310 315 320
 Phe Ser Ala Asn Tyr Gln Asp Arg Ser Tyr Met Met Thr Asn Leu Thr
 325 330 335
 Leu Lys Phe

40

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 959 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

45

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence
 (B) LOCATION: 133...879
 (D) OTHER INFORMATION:

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

	TAAGGAAATG AGTTTTTATA TCATAAAATA AAGTAACCGA GAAAAATCTT TCTCTAAAAA	60
	TAATACTTTT TTAGTTATAA TAACAATTTT GTTTTTTCAA AAACAATAAT TACTATATTT	120
	AGGATTTTAA GA ATG AAT GAC AAG CGT TTT AGA AAA TAT TGT AGT TTT TCT	171
	Met Asn Asp Lys Arg Phe Arg Lys Tyr Cys Ser Phe Ser	
10	1 5 10	
	ATT TTT TTG TCC TTA TTA GGA ACG TTT GAA TTA GAG GCT AAA GAA GAA	219
	Ile Phe Leu Ser Leu Leu Gly Thr Phe Glu Leu Glu Ala Lys Glu Glu	
	15 20 25	
	GAA GAA AAA GAA GAA AGA AAG ACA GAA AGG AAA AAA GAA AAG AAC GCC	267
15	Glu Glu Lys Glu Glu Arg Lys Thr Glu Arg Lys Lys Glu Lys Asn Ala	
	30 35 40 45	
	CAA CAC ACT CTA GGC AAG GTT ACC ACT CAA GCG GCT AAA ATC TTT AAC	315
	Gln His Thr Leu Gly Lys Val Thr Thr Gln Ala Ala Lys Ile Phe Asn	
	50 55 60	
20	TAC AAC AAC CAG ACA ACC ATT TCA AGT AAG GAA TTA GAA AGA AGG CAA	363
	Tyr Asn Asn Gln Thr Thr Ile Ser Ser Lys Glu Leu Glu Arg Arg Gln	
	65 70 75	
	GCC AAC CAA ATC AGC GAC ATG TTT AGA AGA AAC CCT AAT ATC AAT GTG	411
25	Ala Asn Gln Ile Ser Asp Met Phe Arg Arg Asn Pro Asn Ile Asn Val	
	80 85 90	
	GGC GGT GGT GCG GTG ATA GCG CAA AAA ATT TAT GTG CGC GGT ATT GAA	459
	Gly Gly Gly Ala Val Ile Ala Gln Lys Ile Tyr Val Arg Gly Ile Glu	
	95 100 105	
	GAC AGA TTG GCT CGG GTT ACG GTG GAT GGG GCG GCG CAA ATG GGT GCA	507
30	Asp Arg Leu Ala Arg Val Thr Val Asp Gly Ala Ala Gln Met Gly Ala	
	110 115 120 125	
	AGC TAT GGG CAT CAA GGC AAT ACG ATC ATT GAC CCT GGA ATG CTT AAA	555
	Ser Tyr Gly His Gln Gly Asn Thr Ile Ile Asp Pro Gly Met Leu Lys	
	130 135 140	
35	AGC GTG GTG GTT ACT AAA GGG GCG GCT CAA GCG AGC GCG GGG CCT ATG	603
	Ser Val Val Val Thr Lys Gly Ala Ala Gln Ala Ser Ala Gly Pro Met	
	145 150 155	
	GCT TTG ATT GGC GCG ATT AAA ATG GAG ACT AAA AGT GCT AGC GAT TTT	651

Ala Leu Ile Gly Ala Ile Lys Met Glu Thr Lys Ser Ala Ser Asp Phe
 160 165 170

ATC CCT AAA GGT AAA GAC TAC GCC ATA AGT GGG GCT GCC ACT TTT TTA 699
 Ile Pro Lys Gly Lys Asp Tyr Ala Ile Ser Gly Ala Ala Thr Phe Leu

5 175 180 185

ACC AAC TTT GGG GAT CGA GAA ACC GTG ATG GGC GCT TAT CGT CAT AAT 747
 Thr Asn Phe Gly Asp Arg Glu Thr Val Met Gly Ala Tyr Arg His Asn
 190 195 200 205

10 CAT TTT GAT GCG CTT TTG TAT TAC ACG CAT CAA AAT ATT TTT TAC TAT 795
 His Phe Asp Ala Leu Leu Tyr Tyr Thr His Gln Asn Ile Phe Tyr Tyr
 210 215 220

CGT GAT GGG GAT AAT GCT ACA AAA GAT CTC TTT AGA CCT AAA GCG GAG 843
 Arg Asp Gly Asp Asn Ala Thr Lys Asp Leu Phe Arg Pro Lys Ala Glu
 225 230 235

15 AAT AAA GTT ACA GAA GTC CTA GCG AGC AAA ACA ATG TGATGGCTAA GATCAA 895
 Asn Lys Val Thr Glu Val Leu Ala Ser Lys Thr Met
 240 245

TGGTTATTTG AGCGAAAGGG ATATTTTAAC GCTCAGTTAT AACATGACCA GAGACAACGC 955
 TAAC 959

20 (2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 249 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

25 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

30 Met Asn Asp Lys Arg Phe Arg Lys Tyr Cys Ser Phe Ser Ile Phe Leu
 1 5 10 15
 Ser Leu Leu Gly Thr Phe Glu Leu Glu Ala Lys Glu Glu Glu Glu Lys
 20 25 30
 Glu Glu Arg Lys Thr Glu Arg Lys Lys Glu Lys Asn Ala Gln His Thr
 35 40 45
 35 Leu Gly Lys Val Thr Thr Gln Ala Ala Lys Ile Phe Asn Tyr Asn Asn
 50 55 60
 Gln Thr Thr Ile Ser Ser Lys Glu Leu Glu Arg Arg Gln Ala Asn Gln
 65 70 75 80
 Ile Ser Asp Met Phe Arg Arg Asn Pro Asn Ile Asn Val Gly Gly Gly

		85		90		95										
	Ala	Val	Ile	Ala	Gln	Lys	Ile	Tyr	Val	Arg	Gly	Ile	Glu	Asp	Arg	Leu
			100						105					110		
	Ala	Arg	Val	Thr	Val	Asp	Gly	Ala	Ala	Gln	Met	Gly	Ala	Ser	Tyr	Gly
5			115				120					125				
	His	Gln	Gly	Asn	Thr	Ile	Ile	Asp	Pro	Gly	Met	Leu	Lys	Ser	Val	Val
		130				135					140					
	Val	Thr	Lys	Gly	Ala	Ala	Gln	Ala	Ser	Ala	Gly	Pro	Met	Ala	Leu	Ile
	145				150					155					160	
10	Gly	Ala	Ile	Lys	Met	Glu	Thr	Lys	Ser	Ala	Ser	Asp	Phe	Ile	Pro	Lys
				165					170					175		
	Gly	Lys	Asp	Tyr	Ala	Ile	Ser	Gly	Ala	Ala	Thr	Phe	Leu	Thr	Asn	Phe
			180					185					190			
	Gly	Asp	Arg	Glu	Thr	Val	Met	Gly	Ala	Tyr	Arg	His	Asn	His	Phe	Asp
15		195					200					205				
	Ala	Leu	Leu	Tyr	Tyr	Thr	His	Gln	Asn	Ile	Phe	Tyr	Tyr	Arg	Asp	Gly
	210					215					220					
	Asp	Asn	Ala	Thr	Lys	Asp	Leu	Phe	Arg	Pro	Lys	Ala	Glu	Asn	Lys	Val
	225				230					235					240	
20	Thr	Glu	Val	Leu	Ala	Ser	Lys	Thr	Met							
				245												

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1397 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 138...1244
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

	TGAATGCGGG	CATTGGGGCT	AGGTTTGGGC	TTGATTATAA	AGATATTAAT	ATCACCGGAA	60						
35	ATATTGGTAT	GCGCTATGCT	TTTAAATGGT	ATCATTAAAC	CTATTTTAA	CAATCCCAAT	120						
	TCATAGCAGG	ATCACCC	ATG CAA	TTT CAA	AAA GCC	TTA TTA CAT TCA TCA	170						
		Met	Gln	Phe	Gln	Lys	Ala	Leu	Leu	His	Ser	Ser	
		1			5					10			
	TTC TTT	TTA CCT	TTA TTT	TTA TCT	TTT TGT	ATC GCT	GAA GAA	AAT GGG	218				
40	Phe Phe	Leu Pro	Leu Phe	Leu Ser	Phe Cys	Ile Ala	Glu Glu	Asn Gly					
		15			20			25					
	GCG TAT	GCG AGC	GTG GGT	TTT GAA	TAT TCC	ATT AGT	CAT GCC	GTT GAA	266				

Ala Tyr Ala Ser Val Gly Phe Glu Tyr Ser Ile Ser His Ala Val Glu
 30 35 40

CAC AAT AAC CCC TTT TTA AAT CAA GAA CGC ATC CAA ATC ATT TCT AAC 314
 His Asn Asn Pro Phe Leu Asn Gln Glu Arg Ile Gln Ile Ile Ser Asn

5 45 50 55

GCT CAA AAT AAA ATC TAT AAA CTC CAT CAA GTT AAA AAT GAA ATC ACA 362
 Ala Gln Asn Lys Ile Tyr Lys Leu His Gln Val Lys Asn Glu Ile Thr
 60 65 70 75

AGC ATG CCT AAA ACC TTT GCA TAT ATC AAC AAC GCT TTA AAA AAC AAC 410
 10 Ser Met Pro Lys Thr Phe Ala Tyr Ile Asn Asn Ala Leu Lys Asn Asn
 80 85 90

TCC AAA TTA ACC CCC ACT GAA ATG CAA GCC GAA CAA TAC TAC CTC CAA 458
 Ser Lys Leu Thr Pro Thr Glu Met Gln Ala Glu Gln Tyr Tyr Leu Gln
 95 100 105

15 TCC ACC TTT CAA AAC ATT GAA AAA ATA GTA ATG CTT AGC GGT GGC GTT 506
 Ser Thr Phe Gln Asn Ile Glu Lys Ile Val Met Leu Ser Gly Gly Val
 110 115 120

TCA TCT AAC CCA CAA TTA GTC CAA GCG TTG GAA AAA ATG CAA GAA CCC 554
 Ser Ser Asn Pro Gln Leu Val Gln Ala Leu Glu Lys Met Gln Glu Pro
 20 125 130 135

ATT ACT AAC CCT TTA GAA TTT GAA GAA AAC TTA AGA AAT TTA GAA GTG 602
 Ile Thr Asn Pro Leu Glu Phe Glu Glu Asn Leu Arg Asn Leu Glu Val
 140 145 150 155

CAA TTT GCT CAA TCT CAA AAC CGC ATG CTT TCT TCT TTA TCT TCT CAA 650
 25 Gln Phe Ala Gln Ser Gln Asn Arg Met Leu Ser Ser Leu Ser Ser Gln
 160 165 170

ATC GCT GCC ATT TCA AAT TCC TTA AAC GCG CTT GAT CCT AAC TCT TAT 698
 Ile Ala Ala Ile Ser Asn Ser Leu Asn Ala Leu Asp Pro Asn Ser Tyr
 175 180 185

30 TCT AAA AAC ATT TCA AGC ATG TAT GGG GTG AGT TTG AGC GTA GGT TAT 746
 Ser Lys Asn Ile Ser Ser Met Tyr Gly Val Ser Leu Ser Val Gly Tyr
 190 195 200

AAG CAT TTC TTT ACC AAG AAA AAA AAT CAA GGG TTG CGC TAT TAC TTG 794
 Lys His Phe Phe Thr Lys Lys Lys Asn Gln Gly Leu Arg Tyr Tyr Leu
 35 205 210 215

TTT TAT GAC TAT GGT TAC ACT AAT TTT GGT TTT GTG GGC AAT GGC TTT 842
 Phe Tyr Asp Tyr Gly Tyr Thr Asn Phe Gly Phe Val Gly Asn Gly Phe
 220 225 230 235

GAT GGT TTA GGC AAA ATG AAT AAC CAT CTC TAT GGG CTT GGG ATA GAC 890
 Asp Gly Leu Gly Lys Met Asn Asn His Leu Tyr Gly Leu Gly Ile Asp
 240 245 250

5 TAT CTT TAT AAT TTC ATT GAT AAT GCA AAA AAA CAC TCT AGC GTA GGT 938
 Tyr Leu Tyr Asn Phe Ile Asp Asn Ala Lys Lys His Ser Ser Val Gly
 255 260 265

TTT TAT CTG GGT TTT GCT TTA GCG GGG AGT TCG TGG GTA GGG AGT GGT 986
 Phe Tyr Leu Gly Phe Ala Leu Ala Gly Ser Ser Trp Val Gly Ser Gly
 270 275 280

10 TTG AGC ATG TGG GTG AGC CAA ACG GAT TTT ATC AAC AAT TAC TTG ACG 1034
 Leu Ser Met Trp Val Ser Gln Thr Asp Phe Ile Asn Asn Tyr Leu Thr
 285 290 295

GGC TAT CAA GCT AAA ATG CAC ACG AGT TTT TTC CAG ATC CCT TTG AAT 1082
 Gly Tyr Gln Ala Lys Met His Thr Ser Phe Phe Gln Ile Pro Leu Asn
 15 300 305 310 315

TTT GGG GTT CGT GTG AAT GTC AAT AGG CAT AAT GGC TTT GAA ATG GGC 1130
 Phe Gly Val Arg Val Asn Val Asn Arg His Asn Gly Phe Glu Met Gly
 320 325 330

20 TTG AAA ATC CCT TTA GCG ATG AAT TCC TTT TAT GAA ACG CAT GGC AAA 1178
 Leu Lys Ile Pro Leu Ala Met Asn Ser Phe Tyr Glu Thr His Gly Lys
 335 340 345

GGG CTA AAC ACT TCC CTC TTT TTC AAA CGC CTT GTC ATG TTT AAC GTG 1226
 Gly Leu Asn Thr Ser Leu Phe Phe Lys Arg Leu Val Met Phe Asn Val
 350 355 360

25 AGT TAC GTT TAT AGT TTT TAGGGGGGTA GAAATAAGCA CCCCCTTAAA TGTTATCG 1282
 Ser Tyr Val Tyr Ser Phe
 365

CAACCTTTGA ATTTTAAAAA CTCTTTAGTT TTTTGCCTC AAATGATGGA CGCTCTCGCC 1342
 CCCAAGACCA TAATTATTAG AATCGACCTC ATCTATAATG ACCACAATAG AAGCC 1397

30 (2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 369 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

35

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Met Gln Phe Gln Lys Ala Leu Leu His Ser Ser Phe Phe Leu Pro Leu
1 5 10 15
Phe Leu Ser Phe Cys Ile Ala Glu Glu Asn Gly Ala Tyr Ala Ser Val
5 20 25 30
Gly Phe Glu Tyr Ser Ile Ser His Ala Val Glu His Asn Asn Pro Phe
35 40 45
Leu Asn Gln Glu Arg Ile Gln Ile Ile Ser Asn Ala Gln Asn Lys Ile
50 55 60
10 Tyr Lys Leu His Gln Val Lys Asn Glu Ile Thr Ser Met Pro Lys Thr
65 70 75 80
Phe Ala Tyr Ile Asn Asn Ala Leu Lys Asn Asn Ser Lys Leu Thr Pro
85 90 95
15 Thr Glu Met Gln Ala Glu Gln Tyr Tyr Leu Gln Ser Thr Phe Gln Asn
100 105 110
Ile Glu Lys Ile Val Met Leu Ser Gly Gly Val Ser Ser Asn Pro Gln
115 120 125
Leu Val Gln Ala Leu Glu Lys Met Gln Glu Pro Ile Thr Asn Pro Leu
130 135 140
20 Glu Phe Glu Glu Asn Leu Arg Asn Leu Glu Val Gln Phe Ala Gln Ser
145 150 155 160
Gln Asn Arg Met Leu Ser Ser Leu Ser Ser Gln Ile Ala Ala Ile Ser
165 170 175
25 Asn Ser Leu Asn Ala Leu Asp Pro Asn Ser Tyr Ser Lys Asn Ile Ser
180 185 190
Ser Met Tyr Gly Val Ser Leu Ser Val Gly Tyr Lys His Phe Phe Thr
195 200 205
Lys Lys Lys Asn Gln Gly Leu Arg Tyr Tyr Leu Phe Tyr Asp Tyr Gly
210 215 220
30 Tyr Thr Asn Phe Gly Phe Val Gly Asn Gly Phe Asp Gly Leu Gly Lys
225 230 235 240
Met Asn Asn His Leu Tyr Gly Leu Gly Ile Asp Tyr Leu Tyr Asn Phe
245 250 255
35 Ile Asp Asn Ala Lys Lys His Ser Ser Val Gly Phe Tyr Leu Gly Phe
260 265 270
Ala Leu Ala Gly Ser Ser Trp Val Gly Ser Gly Leu Ser Met Trp Val
275 280 285
Ser Gln Thr Asp Phe Ile Asn Asn Tyr Leu Thr Gly Tyr Gln Ala Lys
290 295 300
40 Met His Thr Ser Phe Phe Gln Ile Pro Leu Asn Phe Gly Val Arg Val
305 310 315 320
Asn Val Asn Arg His Asn Gly Phe Glu Met Gly Leu Lys Ile Pro Leu
325 330 335
45 Ala Met Asn Ser Phe Tyr Glu Thr His Gly Lys Gly Leu Asn Thr Ser
340 345 350
Leu Phe Phe Lys Arg Leu Val Met Phe Asn Val Ser Tyr Val Tyr Ser
355 360 365
Phe

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1030 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence

(B) LOCATION: 342...824

(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

CACTCTAAGC	GTCAAAC TCT	CTTTTTCTTT	AGAGGAAGAA	AGCAAGCGGA	TCCATCTTAA	60
AGCCTTACAA	AATATCTTAA	ATAACGCTAA	AAGCGCGCAT	TTTAAATTTG	TTTTAGAGAG	120
CCAAAACGCC	GCTCAATCTA	TTATAGAAAT	TCAAAGCCTC	TTGAAACAAC	TCTCCTTAAA	180
AAATAATGAA	ATCTTTTTTA	TGCCTTTAGG	CACAAATAAC	AACGAGCTAG	ACAAAAATCT	240
AAAAACCCTA	GCCCCCCTAG	CCATAAAGCA	TGGTTTCAGG	CTAAGCGATA	GGCTTCATAT	300
CCGCTTG TGG	GATAATCAAA	AAGGGTTTTA	AAAAGTTAAT	C ATG ACC ATC AAA GTT		356
				Met Thr Ile Lys Val		

TTT	TCG	CCC	AAA	TAC	CCC	ACT	GAA	TTA	GAA	GAA	TTT	TAT	GCT	GAG	CGT	404
Phe	Ser	Pro	Lys	Tyr	Pro	Thr	Glu	Leu	Glu	Glu	Phe	Tyr	Ala	Glu	Arg	
				10				15					20			

ATC GCT GAC AAC CCT TTA GGG TTT ATC CAA CGC TTG GAT CTT TTG CCT 452
Ile Ala Asp Asn Pro Leu Gly Phe Ile Gln Arg Leu Asp Leu Leu Pro
25 30 35

AGT ATT AGC GGG TTC GTT CAA AAA TTG CGC GAG CAT GGC GGG GAA TTT 500
Ser Ile Ser Gly Phe Val Gln Lys Leu Arg Glu His Gly Gly Glu Phe
40 45 50

TTT GAA ATG AGA GAG GGT AAC AAG CTC ATT GGG ATT TGT GGG CTT AAT 548
Phe Glu Met Arg Glu Gly Asn Lys Leu Ile Gly Ile Cys Gly Leu Asn
55 60 65

CCT ATC AAT CAA ACA GAA GCC GAG CTG TGC AAA TTC CAC ATA AAT AGT 596
Pro Ile Asn Gln Thr Glu Ala Glu Leu Cys Lys Phe His Ile Asn Ser
70 75 80 85

GCT TAT CAA TCC CAA GGG CTA GGT CAA AAA CTC TAT GAG AGC GTG GAG 644
Ala Tyr Gln Ser Gln Gly Leu Gly Gln Lys Leu Tyr Glu Ser Val Glu
90 95 100

	AAA TAC GCT TTC ATT AAA GGC TAT ACT AAA ATC TCT CTG CAT GTG AGC	692
	Lys Tyr Ala Phe Ile Lys Gly Tyr Thr Lys Ile Ser Leu His Val Ser	
	105 110 115	
	AAA AGC CAA ATC AAG GCA TGC AAC CTC TAT CAA AAG CTG GGT TTT GTG	740
5	Lys Ser Gln Ile Lys Ala Cys Asn Leu Tyr Gln Lys Leu Gly Phe Val	
	120 125 130	
	CAC ATC AAA GAA GAG GAT TGC GTG GTG GAG TTG GGC GAA GAG ACT TTG	788
	His Ile Lys Glu Glu Asp Cys Val Val Glu Leu Gly Glu Glu Thr Leu	
	135 140 145	
10	ATT TTC CCC ACT CTT TTT ATG GAA AAG ATT TTG TCT TGATTGGTGC ATCCAT	840
	Ile Phe Pro Thr Leu Phe Met Glu Lys Ile Leu Ser	
	150 155 160	
	TTGACACACG CCCAAGCGAC ATTCAAATA TCAAACCTTTC ATTAACACAA CCCAATTAAC	900
	GCTAAATAAA CCCTAAACA AACACTCGTT GTTAAATTT TGTTTTTCAA GCGCTTCGCA	960
15	AAGTTTTAGA AGCCCTATTT AGGGGTAAAC GCTAAAATAG GCTATCAAAA CTACTTTAAT	1020
	GATTTTATAG	1030

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 161 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

	Met Thr Ile Lys Val Phe Ser Pro Lys Tyr Pro Thr Glu Leu Glu Glu	
	1 5 10 15	
	Phe Tyr Ala Glu Arg Ile Ala Asp Asn Pro Leu Gly Phe Ile Gln Arg	
	20 25 30	
30	Leu Asp Leu Leu Pro Ser Ile Ser Gly Phe Val Gln Lys Leu Arg Glu	
	35 40 45	
	His Gly Gly Glu Phe Phe Glu Met Arg Glu Gly Asn Lys Leu Ile Gly	
	50 55 60	
	Ile Cys Gly Leu Asn Pro Ile Asn Gln Thr Glu Ala Glu Leu Cys Lys	
35	65 70 75 80	
	Phe His Ile Asn Ser Ala Tyr Gln Ser Gln Gly Leu Gly Gln Lys Leu	
	85 90 95	
	Tyr Glu Ser Val Glu Lys Tyr Ala Phe Ile Lys Gly Tyr Thr Lys Ile	
	100 105 110	
40	Ser Leu His Val Ser Lys Ser Gln Ile Lys Ala Cys Asn Leu Tyr Gln	
	115 120 125	

Lys Leu Gly Phe Val His Ile Lys Glu Glu Asp Cys Val Val Glu Leu
130 135 140
Gly Glu Glu Thr Leu Ile Phe Pro Thr Leu Phe Met Glu Lys Ile Leu
145 150 155 160
5 Ser

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 1477 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:

- 15 (A) NAME/KEY: Coding Sequence
(B) LOCATION: 374...1267
(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

CGTGGAGTTT TTTAGGCATT TCTTTATATT CATTCAATAA CGCTTGCGCG GGCAATTCTT 60
20 CAACTAAAAT CTCTACTAAC AATTCATCTG AATGCAAAAT CTCAATTCTC CCTAAAAAAC 120
AAAATCACTT TTAAGACTAA ATCATGTTAG AATTATACTT GAATTTACAC TCAGTTTAGT 180
TTATTTCTTA ATACAAAAGG TAGGCGTTTT GAAACATTTA ACCCCACTCA CTCACACCAT 240
CTTTAAAGCC TTATGGCTAG GCACAGCCTT AAGTGCATCT TTAAGTTTAG CCGCAACAGA 300
AAGCCCCACT AAAACAGAGC CTAAGCCCCG TAAAGGGGTT AAAACAAGC CCAAATCGCC 360
25 CGTTACTAAA GTC ATG ATG ACC AAT TGC GAC AAT ATT AAA GAT TTT AAC 409
Met Met Thr Asn Cys Asp Asn Ile Lys Asp Phe Asn
1 5 10
GCT AAG CAA AAA GAA GTC TTA AAA GCC GCT TAT CAA TTC GGC TCT AAA 457
30 Ala Lys Gln Lys Glu Val Leu Lys Ala Ala Tyr Gln Phe Gly Ser Lys
15 20 25
GAA AAT TTA GGC TAT GAA ATG GCA GGC ATT GCA TGG AAA GAG TCA TGC 505
Glu Asn Leu Gly Tyr Glu Met Ala Gly Ile Ala Trp Lys Glu Ser Cys
30 35 40
GCA GGG GTT TAT AAA ATC AAT TTT TCG GAT CCG AGC GCG GGC GTG TAT 553
35 Ala Gly Val Tyr Lys Ile Asn Phe Ser Asp Pro Ser Ala Gly Val Tyr
45 50 55 60
CAT TCT TAT ATC CCA AGC GTT CTA AAA AGC TAT GGG CAT AAT GAT AGC 601
His Ser Tyr Ile Pro Ser Val Leu Lys Ser Tyr Gly His Asn Asp Ser
65 70 75

	CCC TTT TTG CGT AAT GTG ATG GGG GAA TTG CTC ATT AAA GAC GAT GCG	649
	Pro Phe Leu Arg Asn Val Met Gly Glu Leu Leu Ile Lys Asp Asp Ala	
	80 85 90	
5	TTT GCT TCT GAA GTG GCT TTA AAA GAG TTG CTC TAT TGG AAA ACA CGC	697
	Phe Ala Ser Glu Val Ala Leu Lys Glu Leu Leu Tyr Trp Lys Thr Arg	
	95 100 105	
	TAC CAT GAC AAT TTA AAA GAC ATG ATT AAA TCT TAC AAC AAG GGC AGT	745
	Tyr His Asp Asn Leu Lys Asp Met Ile Lys Ser Tyr Asn Lys Gly Ser	
	110 115 120	
10	CGT TGG GAA AGG AGC GAA AAA TCT AAC GCT GAT GCT GAA AAA TAT TAC	793
	Arg Trp Glu Arg Ser Glu Lys Ser Asn Ala Asp Ala Glu Lys Tyr Tyr	
	125 130 135 140	
15	GAA GAG ATA CAA GAC AGA ATC AGG CGT TTG AAA GAA TCT AAA ATC TTT	841
	Glu Glu Ile Gln Asp Arg Ile Arg Arg Leu Lys Glu Ser Lys Ile Phe	
	145 150 155	
	GAT TCG CAG TCT AGT AAT GAC CAA GAA TTG CAA AAA AGC GCT AAT AGC	889
	Asp Ser Gln Ser Ser Asn Asp Gln Glu Leu Gln Lys Ser Ala Asn Ser	
	160 165 170	
20	AAC CTG GAT TTA GAC CCT ATC GGC AAC GCC ATG CCC CAA GCC TTA ATT	937
	Asn Leu Asp Leu Asp Pro Ile Gly Asn Ala Met Pro Gln Ala Leu Ile	
	175 180 185	
	GCC AAA GAA ACT AAA ATA GAA GAA ACC CAA GCA GAA AAA TCC CAA GAA	985
	Ala Lys Glu Thr Lys Ile Glu Glu Thr Gln Ala Glu Lys Ser Gln Glu	
	190 195 200	
25	ATG AAA GAG ACA ACT AGC GAG CAA ACA AAA AGT AAG CCA GAA AAA GCA	1033
	Met Lys Glu Thr Thr Ser Glu Gln Thr Lys Ser Lys Pro Glu Lys Ala	
	205 210 215 220	
30	AAA GAT AAA CCC ATG TAT TTG GCG CAA ATC AAC AGC ACT GAT TTC ACA	1081
	Lys Asp Lys Pro Met Tyr Leu Ala Gln Ile Asn Ser Thr Asp Phe Thr	
	225 230 235	
	CCC GTT AAA AAA AGC CCC AAA AAA CCG GCT AAA GTG AGC CAA AAA CAC	1129
	Pro Val Lys Lys Ser Pro Lys Lys Pro Ala Lys Val Ser Gln Lys His	
	240 245 250	
35	TCC TTT AAG AAT AAC ATT AAA AAT AAT GTA AAA AAC AAC GCC AAA ACC	1177
	Ser Phe Lys Asn Asn Ile Lys Asn Asn Val Lys Asn Asn Ala Lys Thr	
	255 260 265	
	GCT TCC AAA AAA CAA GAA ATG TGC AAA AAT TGC TCT CCA GGG CAA AGG	1225
	Ala Ser Lys Lys Gln Glu Met Cys Lys Asn Cys Ser Pro Gly Gln Arg	

270

275

280

AAT GCG ATT TTA GCT AAC CAC ATC ACT CTC ATG CAA GAG CTT TAAAAAGTC 1276
 Asn Ala Ile Leu Ala Asn His Ile Thr Leu Met Gln Glu Leu
 285 290 295

5 CTAAAAATGG CGCAAAAAAC TCTTTTGATT ATCACTGATG GCATTGGGTA TCGTAAAGAT 1336
 AGCGATCATA ACGCTTTCTT CCATGCCAAA AAACCCACTT ATGATTTGAT GTTTAAAACC 1396
 TTGCCTTATA GCCTGATTGA TACGCATGGC TTGAGCGTGG GCTTACCTAA GGGGCAAATG 1456
 GGAAATTCTG AAGTGGGGCA T 1477

(2) INFORMATION FOR SEQ ID NO:26:

10 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 298 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

15 (ii) MOLECULE TYPE: protein
 (v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

	Met	Met	Thr	Asn	Cys	Asp	Asn	Ile	Lys	Asp	Phe	Asn	Ala	Lys	Gln	Lys	
	1				5					10					15		
20	Glu	Val	Leu	Lys	Ala	Ala	Tyr	Gln	Phe	Gly	Ser	Lys	Glu	Asn	Leu	Gly	
				20					25					30			
	Tyr	Glu	Met	Ala	Gly	Ile	Ala	Trp	Lys	Glu	Ser	Cys	Ala	Gly	Val	Tyr	
			35				40						45				
25	Lys	Ile	Asn	Phe	Ser	Asp	Pro	Ser	Ala	Gly	Val	Tyr	His	Ser	Tyr	Ile	
		50				55						60					
	Pro	Ser	Val	Leu	Lys	Ser	Tyr	Gly	His	Asn	Asp	Ser	Pro	Phe	Leu	Arg	
	65				70					75					80		
	Asn	Val	Met	Gly	Glu	Leu	Leu	Ile	Lys	Asp	Asp	Ala	Phe	Ala	Ser	Glu	
				85					90						95		
30	Val	Ala	Leu	Lys	Glu	Leu	Leu	Tyr	Trp	Lys	Thr	Arg	Tyr	His	Asp	Asn	
				100					105					110			
	Leu	Lys	Asp	Met	Ile	Lys	Ser	Tyr	Asn	Lys	Gly	Ser	Arg	Trp	Glu	Arg	
			115				120						125				
	Ser	Glu	Lys	Ser	Asn	Ala	Asp	Ala	Glu	Lys	Tyr	Tyr	Glu	Glu	Ile	Gln	
35		130				135						140					
	Asp	Arg	Ile	Arg	Arg	Leu	Lys	Glu	Ser	Lys	Ile	Phe	Asp	Ser	Gln	Ser	
	145				150					155					160		
	Ser	Asn	Asp	Gln	Glu	Leu	Gln	Lys	Ser	Ala	Asn	Ser	Asn	Leu	Asp	Leu	
				165					170					175			
40	Asp	Pro	Ile	Gly	Asn	Ala	Met	Pro	Gln	Ala	Leu	Ile	Ala	Lys	Glu	Thr	
				180					185					190			
	Lys	Ile	Glu	Glu	Thr	Gln	Ala	Glu	Lys	Ser	Gln	Glu	Met	Lys	Glu	Thr	
		195					200						205				

	Thr	Ser	Glu	Gln	Thr	Lys	Ser	Lys	Pro	Glu	Lys	Ala	Lys	Asp	Lys	Pro
	210					215					220					
	Met	Tyr	Leu	Ala	Gln	Ile	Asn	Ser	Thr	Asp	Phe	Thr	Pro	Val	Lys	Lys
	225				230					235					240	
5	Ser	Pro	Lys	Lys	Pro	Ala	Lys	Val	Ser	Gln	Lys	His	Ser	Phe	Lys	Asn
				245					250					255		
	Asn	Ile	Lys	Asn	Asn	Val	Lys	Asn	Asn	Ala	Lys	Thr	Ala	Ser	Lys	Lys
			260				265					270				
	Gln	Glu	Met	Cys	Lys	Asn	Cys	Ser	Pro	Gly	Gln	Arg	Asn	Ala	Ile	Leu
10		275				280						285				
	Ala	Asn	His	Ile	Thr	Leu	Met	Gln	Glu	Leu						
	290					295										

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- 15 (A) LENGTH: 1515 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:

- 20 (A) NAME/KEY: Coding Sequence
(B) LOCATION: 141...1340
(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

25	TTAGTGTGGA	TTTTTTTATC	GTTAGTGTTC	GTGCGTCCTT	TAGAGGCTTT	GAGCGTGTTC	60
	ATGGGGTTGT	ATTTGATTTA	TGGCATCATT	CGGTGGCTCT	TTTAAATGGT	AAAAATTATT	120
	TTTAATAAAA	ATAAAAGCGC	ATG AAA GAA	TCT TTT TAC	ATA GAG GGA	ATG	170
			Met Lys Glu	Ser Phe Tyr	Ile Glu Gly	Met	
			1	5	10		
30	ACT TGC ACG	GCG TGT TCT	AGC GGG ATT	GAA CGC TCT	TTG GGG CGT	AAG	218
	Thr Cys Thr	Ala Cys Ser	Ser Gly Ile	Glu Arg Ser	Leu Gly Arg	Lys	
		15	20	25			
	AGT TTT GTG	AAA AAA ATA	GAA GTG AGC	CTT TTA AAT	AAG AGC GCT	AAC	266
	Ser Phe Val	Lys Lys Ile	Glu Val Ser	Leu Leu Asn	Lys Ser Ala	Asn	
35		30	35	40			
	ATT GAA TTT	GAC GAA AAC	CAA ACC AAT	TTA GAC GAA	ATT TTT AAA	CTC	314
	Ile Glu Phe	Asp Glu Asn	Gln Thr Asn	Leu Asp Glu	Ile Phe Lys	Leu	
	45	50	55				
	ATT GAA AAG	CTA GGC TAT	AGC CCT AAA	AAA GCT CTG	ACA AAA GAA	AAA	362
40	Ile Glu Lys	Leu Gly Tyr	Ser Pro Lys	Lys Ala Leu	Thr Lys Glu	Lys	

	60	65	70	
	AAA GAA TTT TTT AGC CCT AAT GTT AAA TTA GCG TTA GCG GTT ATT TTC			410
	Lys Glu Phe Phe Ser Pro Asn Val Lys Leu Ala Leu Ala Val Ile Phe			
	75	80	85	90
5	ACG CTT TTT GTG GTG TAT CTT TCT ATG GGG GCG ATG CTT AGC CCT AGC			458
	Thr Leu Phe Val Val Tyr Leu Ser Met Gly Ala Met Leu Ser Pro Ser			
		95	100	105
	CTT TTA CCT GAA AGC TTG CTT GCA ATT GAT AAT CAT AGT AAT TTT TTA			506
10	Leu Leu Pro Glu Ser Leu Leu Ala Ile Asp Asn His Ser Asn Phe Leu			
		110	115	120
	AAC GCT TGC TTA CAG CTT ATA GGC GCA CTC ATT GTC ATG CAT TTG GGG			554
	Asn Ala Cys Leu Gln Leu Ile Gly Ala Leu Ile Val Met His Leu Gly			
		125	130	135
15	AGG GAT TTT TAC ATT CAA GGG TTT AAA GCC TTA TGG CAC AGA CAA CCC			602
	Arg Asp Phe Tyr Ile Gln Gly Phe Lys Ala Leu Trp His Arg Gln Pro			
		140	145	150
	AAC ATG AGC AGC CTT ATC GCC ATA GGC ACA AGC GCT GCC TTA ATT TCA			650
	Asn Met Ser Ser Leu Ile Ala Ile Gly Thr Ser Ala Ala Leu Ile Ser			
		155	160	165
20	AGC CTG TGG CAA TTG TAT TTG GTC TAT ACC AAT CAT TAT ACC GAT CAG			698
	Ser Leu Trp Gln Leu Tyr Leu Val Tyr Thr Asn His Tyr Thr Asp Gln			
		175	180	185
	TGG TCT TAT GGG CAT TAT TAT TTT GAA AGC GTG TGC GTG ATT TTA ATG			746
25	Trp Ser Tyr Gly His Tyr Tyr Phe Glu Ser Val Cys Val Ile Leu Met			
		190	195	200
	TTT GTG ATG GTG GGC AAA CGC ATT GAA AAT GTT TCT AAA GAC AAA GCT			794
	Phe Val Met Val Gly Lys Arg Ile Glu Asn Val Ser Lys Asp Lys Ala			
		205	210	215
30	TTA GAC GCT ATG CAA GCC TTG ATG AAA AAC GCC CCA AAA ACC GCC CTT			842
	Leu Asp Ala Met Gln Ala Leu Met Lys Asn Ala Pro Lys Thr Ala Leu			
		220	225	230
	AAA ATG CAA AAT AAC CAA CAG ATT GAA GTT TTA GTG GAT AGC ATT GTG			890
	Lys Met Gln Asn Asn Gln Gln Ile Glu Val Leu Val Asp Ser Ile Val			
		235	240	245
35	GTG GGG GAT ATT CTA AAA GTC CTC CCT GGA AGC GCG ATT GCG GTG GAT			938
	Val Gly Asp Ile Leu Lys Val Leu Pro Gly Ser Ala Ile Ala Val Asp			
		255	260	265

	GGT GAA ATC ATA GAG GGC GAA GGG GAA TTA GAT GAG AGC ATG TTG AGC	986
	Gly Glu Ile Ile Glu Gly Glu Gly Glu Leu Asp Glu Ser Met Leu Ser	
	270 275 280	
	GGC GAA GCG TTG CCG GTT TAT AAA AAA GTC GGC GAT AAA GTC TTT TCA	1034
5	Gly Glu Ala Leu Pro Val Tyr Lys Lys Val Gly Asp Lys Val Phe Ser	
	285 290 295	
	GGG ACA TTC AAT AGC CAC ACG AGT TTT TTA ATG AAA GCC ACG CAA AAC	1082
	Gly Thr Phe Asn Ser His Thr Ser Phe Leu Met Lys Ala Thr Gln Asn	
	300 305 310	
10	AAC AAA AAC AGC ACC TTG TCT CAA ATT ATA GAA ATG ATT TAT AAC GCT	1130
	Asn Lys Asn Ser Thr Leu Ser Gln Ile Ile Glu Met Ile Tyr Asn Ala	
	315 320 325 330	
	CAA AGT TCA AAG GCA GAG ATT TCT CGC TTA GCG GAT AAG GTT TCA AGC	1178
	Gln Ser Ser Lys Ala Glu Ile Ser Arg Leu Ala Asp Lys Val Ser Ser	
15	335 340 345	
	GTG TTT GTG CCA AGC GTG ATC GCT ATT TCT ATT TTA GCG TTT GTG GTG	1226
	Val Phe Val Pro Ser Val Ile Ala Ile Ser Ile Leu Ala Phe Val Val	
	350 355 360	
	TGG CTC ATC ATT GCA CCT AAG CCC GAT TTT TGG TGG AAT TTT GGA ATC	1274
20	Trp Leu Ile Ile Ala Pro Lys Pro Asp Phe Trp Trp Asn Phe Gly Ile	
	365 370 375	
	GCT TTA GAA GTG TTT GTA TCG GTT TTA GTG ATT TCT TGC CCT TGC GCT	1322
	Ala Leu Glu Val Phe Val Ser Val Leu Val Ile Ser Cys Pro Cys Ala	
	380 385 390	
25	TTA GGA TTG CTA CGC CTA TGAGCATTTT AGTAGCGAAC CAGAAAGCGA GTTCTTTA	1378
	Leu Gly Leu Leu Arg Leu	
	395 400	
	GGGTTATTTT TTAAAGACGC TAAAAGTTTA GAAAAAGCAA GGCTAGTCAA TACGATCGTT	1438
	TTTGATAAAA CCGGCACGCT CACTAACGGC AAGCCTGTCTG TTAAAAGCGT TCATTCTAAG	1498
30	ATAGAATTAT TAGAGTT	1515

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

- 35 (A) LENGTH: 400 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Met Lys Glu Ser Phe Tyr Ile Glu Gly Met Thr Cys Thr Ala Cys Ser
1 5 10 15
Ser Gly Ile Glu Arg Ser Leu Gly Arg Lys Ser Phe Val Lys Lys Ile
5 20 25 30
Glu Val Ser Leu Leu Asn Lys Ser Ala Asn Ile Glu Phe Asp Glu Asn
35 40 45
Gln Thr Asn Leu Asp Glu Ile Phe Lys Leu Ile Glu Lys Leu Gly Tyr
50 55 60
10 Ser Pro Lys Lys Ala Leu Thr Lys Glu Lys Lys Glu Phe Phe Ser Pro
65 70 75 80
Asn Val Lys Leu Ala Leu Ala Val Ile Phe Thr Leu Phe Val Val Tyr
85 90 95
Leu Ser Met Gly Ala Met Leu Ser Pro Ser Leu Leu Pro Glu Ser Leu
15 100 105 110
Leu Ala Ile Asp Asn His Ser Asn Phe Leu Asn Ala Cys Leu Gln Leu
115 120 125
Ile Gly Ala Leu Ile Val Met His Leu Gly Arg Asp Phe Tyr Ile Gln
130 135 140
20 Gly Phe Lys Ala Leu Trp His Arg Gln Pro Asn Met Ser Ser Leu Ile
145 150 155 160
Ala Ile Gly Thr Ser Ala Ala Leu Ile Ser Ser Leu Trp Gln Leu Tyr
165 170 175
Leu Val Tyr Thr Asn His Tyr Thr Asp Gln Trp Ser Tyr Gly His Tyr
25 180 185 190
Tyr Phe Glu Ser Val Cys Val Ile Leu Met Phe Val Met Val Gly Lys
195 200 205
Arg Ile Glu Asn Val Ser Lys Asp Lys Ala Leu Asp Ala Met Gln Ala
210 215 220
30 Leu Met Lys Asn Ala Pro Lys Thr Ala Leu Lys Met Gln Asn Asn Gln
225 230 235 240
Gln Ile Glu Val Leu Val Asp Ser Ile Val Val Gly Asp Ile Leu Lys
245 250 255
Val Leu Pro Gly Ser Ala Ile Ala Val Asp Gly Glu Ile Ile Glu Gly
35 260 265 270
Glu Gly Glu Leu Asp Glu Ser Met Leu Ser Gly Glu Ala Leu Pro Val
275 280 285
Tyr Lys Lys Val Gly Asp Lys Val Phe Ser Gly Thr Phe Asn Ser His
290 295 300
40 Thr Ser Phe Leu Met Lys Ala Thr Gln Asn Asn Lys Asn Ser Thr Leu
305 310 315 320
Ser Gln Ile Ile Glu Met Ile Tyr Asn Ala Gln Ser Ser Lys Ala Glu
325 330 335
Ile Ser Arg Leu Ala Asp Lys Val Ser Ser Val Phe Val Pro Ser Val
45 340 345 350
Ile Ala Ile Ser Ile Leu Ala Phe Val Val Trp Leu Ile Ile Ala Pro
355 360 365
Lys Pro Asp Phe Trp Trp Asn Phe Gly Ile Ala Leu Glu Val Phe Val
370 375 380

Ser Val Leu Val Ile Ser Cys Pro Cys Ala Leu Gly Leu Leu Arg Leu
 385 390 395 400

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 1443 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ix) FEATURE:

- 10 (A) NAME/KEY: Coding Sequence
 (B) LOCATION: 76...1389
 (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

15 ACTTTAAAAA ACCCCCTTAA AAAGGTTTTT AGGTATAATT AGCGATCTTT TAGTTTCAAA 60
 TAGTAGAGAG ATGGG ATG AAA AAA ATA TGG CTT TTA GTG TGG GGC TTG TGT 111
 Met Lys Lys Ile Trp Leu Leu Val Trp Gly Leu Cys
 1 5 10

20 TCT TGG GTG TTT TTG CAT GCG ATA GAG ATG ATA GAA AAA GCC CCT ACA 159
 Ser Trp Val Phe Leu His Ala Ile Glu Met Ile Glu Lys Ala Pro Thr
 15 20 25

AAT GTA GAG GAT AGA GAC AAA GCC CCC CAT TTG TTG CTT TTA GCA GGG 207
 Asn Val Glu Asp Arg Asp Lys Ala Pro His Leu Leu Leu Ala Gly
 30 35 40

25 ATT CAA GGC GAT GAG CCT GGT GGG TTT AAT GCA ACT AAT TTG TTT TTA 255
 Ile Gln Gly Asp Glu Pro Gly Gly Phe Asn Ala Thr Asn Leu Phe Leu
 45 50 55 60

ATG CAT TAT AGC GTT TTA AAA GGT TTG GTT GAA GTG GTT CCT GTA TTG 303
 Met His Tyr Ser Val Leu Lys Gly Leu Val Glu Val Val Pro Val Leu
 30 65 70 75

AAT AAG CCT TCC ATG TTA AGA AAT CAT AGG GGC TTG TAT GGG GAT ATG 351
 Asn Lys Pro Ser Met Leu Arg Asn His Arg Gly Leu Tyr Gly Asp Met
 80 85 90

35 AAC CGC AAA TTT GCC GCT TTA GAC AAG AAT GAC CCT GAA TAC CCC ACT 399
 Asn Arg Lys Phe Ala Ala Leu Asp Lys Asn Asp Pro Glu Tyr Pro Thr
 95 100 105

ATC CAG GAA ATC AAA TCC TTG ATT GCA AAA CCC AGT ATA GAC GCT GTC 447

	Ile	Gln	Glu	Ile	Lys	Ser	Leu	Ile	Ala	Lys	Pro	Ser	Ile	Asp	Ala	Val	
	110						115					120					
	TTG	CAT	TTG	CAT	GAT	GGC	GGT	GGG	TAT	TAC	CGC	CCT	GTT	TAT	GTT	GAT	495
	Leu	His	Leu	His	Asp	Gly	Gly	Gly	Tyr	Tyr	Arg	Pro	Val	Tyr	Val	Asp	
5	125					130					135					140	
	GCG	ATG	CTC	AAT	CCT	AAG	CGC	TGG	GGG	AAT	TGC	TTT	ATT	ATT	GAT	CAA	543
	Ala	Met	Leu	Asn	Pro	Lys	Arg	Trp	Gly	Asn	Cys	Phe	Ile	Ile	Asp	Gln	
					145					150					155		
	GAT	GAG	GTT	AAA	GGG	GCG	AAA	TTC	CCT	AAT	TTG	CTT	GCT	TTT	GCA	AAC	591
10	Asp	Glu	Val	Lys	Gly	Ala	Lys	Phe	Pro	Asn	Leu	Leu	Ala	Phe	Ala	Asn	
				160					165					170			
	AAT	ACG	ATT	GAG	AGT	ATC	AAC	GCC	CAT	TTA	TTG	CAC	CCC	ATT	GAA	GAG	639
	Asn	Thr	Ile	Glu	Ser	Ile	Asn	Ala	His	Leu	Leu	His	Pro	Ile	Glu	Glu	
				175				180					185				
15	TAT	CAT	TTA	AAA	AAC	ACG	CGC	ACC	GCG	CAA	GGC	GAT	ACA	GAA	ATG	CAA	687
	Tyr	His	Leu	Lys	Asn	Thr	Arg	Thr	Ala	Gln	Gly	Asp	Thr	Glu	Met	Gln	
				190				195				200					
	AAA	GCC	CTA	ACT	TTT	TAT	GCG	ATC	AAC	CAA	AAA	AAG	AGC	GCT	TTT	GCC	735
	Lys	Ala	Leu	Thr	Phe	Tyr	Ala	Ile	Asn	Gln	Lys	Lys	Ser	Ala	Phe	Ala	
20	205					210					215				220		
	AAT	GAA	GCT	AGC	AAA	GAA	CTC	CCT	TTA	GCA	TCA	AGG	GTG	TTT	TAC	CAC	783
	Asn	Glu	Ala	Ser	Lys	Glu	Leu	Pro	Leu	Ala	Ser	Arg	Val	Phe	Tyr	His	
					225					230				235			
	CTG	CAA	GCC	ATT	GAG	GGC	TTA	CTC	AAT	CAG	CTC	AAT	ATC	CCT	TTT	AAG	831
25	Leu	Gln	Ala	Ile	Glu	Gly	Leu	Leu	Asn	Gln	Leu	Asn	Ile	Pro	Phe	Lys	
				240					245					250			
	CGC	GAT	TTT	GAT	CTT	AAC	CCT	AAC	AGC	GTG	CAT	GCC	CTA	ATC	AAT	GAT	879
	Arg	Asp	Phe	Asp	Leu	Asn	Pro	Asn	Ser	Val	His	Ala	Leu	Ile	Asn	Asp	
				255				260					265				
30	AAA	AAC	TTG	TGG	GCA	AAA	ATC	AGC	TCT	TTG	CCT	AAA	ATG	CCC	CTT	TTT	927
	Lys	Asn	Leu	Trp	Ala	Lys	Ile	Ser	Ser	Leu	Pro	Lys	Met	Pro	Leu	Phe	
				270			275					280					
	AAC	TTG	CGC	CCT	AAA	CTC	AAT	CAT	TTC	CCC	TTA	CCC	CAC	AAC	ACT	AAA	975
	Asn	Leu	Arg	Pro	Lys	Leu	Asn	His	Phe	Pro	Leu	Pro	His	Asn	Thr	Lys	
35	285					290					295					300	
	ATC	CCA	CAA	ATC	CCC	ATA	GAG	AGC	AAC	GCT	TAC	ATT	GTA	GGG	CTA	GTC	1023
	Ile	Pro	Gln	Ile	Pro	Ile	Glu	Ser	Asn	Ala	Tyr	Ile	Val	Gly	Leu	Val	
				305					310					315			

AAA AAT AAA CAA GAA GTG TTT TTA AAA TAC GGC AAC AAG CTC ATG ACA 1071
 Lys Asn Lys Gln Glu Val Phe Leu Lys Tyr Gly Asn Lys Leu Met Thr
 320 325 330

CGA TTA TCG CCT TTT TAC ATA GAG TTT GAT CCT TCT TTA GAA GAA GTG 1119
 5 Arg Leu Ser Pro Phe Tyr Ile Glu Phe Asp Pro Ser Leu Glu Glu Val
 335 340 345

AAA ATG CAA ATT GAC AAT AAG GAT CAA ATG GTT AAA ATA GGG AGC GTG 1167
 Lys Met Gln Ile Asp Asn Lys Asp Gln Met Val Lys Ile Gly Ser Val
 350 355 360

10 GTT GAA GTG AAA GAG AGT TTT TAT ATC CAT GCT ATG GAC AAT ATC CGT 1215
 Val Glu Val Lys Glu Ser Phe Tyr Ile His Ala Met Asp Asn Ile Arg
 365 370 375 380

GCG AAT GTG ATT GGC TTT AGC GTT TCT AAT GAA AAT AAG CCT AAT GAA 1263
 15 Ala Asn Val Ile Gly Phe Ser Val Ser Asn Glu Asn Lys Pro Asn Glu
 385 390 395

GCG GGT TAT ACG ATT AAA TTT AAA GAT TTT CAA AAA CGC TTT TCA TTG 1311
 Ala Gly Tyr Thr Ile Lys Phe Lys Asp Phe Gln Lys Arg Phe Ser Leu
 400 405 410

GAC AAG CAA GAA AGG ATC TAT CGC ATA GAA TTT TAT AAA AAC AAC GCG 1359
 20 Asp Lys Gln Glu Arg Ile Tyr Arg Ile Glu Phe Tyr Lys Asn Asn Ala
 415 420 425

TTT AGC GGG ATG ATC TTA GTG AAA TTT GTG TAGGAATGGA TAAATCTCAT TGC 1412
 Phe Ser Gly Met Ile Leu Val Lys Phe Val
 430 435

25 CTTTTAACAT TCAAGGGTTT TGGTATTTTT T 1443

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 438 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

35 Met Lys Lys Ile Trp Leu Leu Val Trp Gly Leu Cys Ser Trp Val Phe
 1 5 10 15
 Leu His Ala Ile Glu Met Ile Glu Lys Ala Pro Thr Asn Val Glu Asp

		20		25		30	
	Arg	Asp	Lys	Ala	Pro	His	Leu
		35		40		45	
5	Glu	Pro	Gly	Gly	Phe	Asn	Ala
	50		55		60		
	Val	Leu	Lys	Gly	Leu	Val	Glu
	65		70		75		80
	Met	Leu	Arg	Asn	His	Arg	Gly
		85		90		95	
10	Ala	Ala	Leu	Asp	Lys	Asn	Asp
		100		105		110	
	Lys	Ser	Leu	Ile	Ala	Lys	Pro
		115		120		125	
	Asp	Gly	Gly	Gly	Tyr	Tyr	Arg
15		130		135		140	
	Pro	Lys	Arg	Trp	Gly	Asn	Cys
	145		150		155		160
	Gly	Ala	Lys	Phe	Pro	Asn	Leu
		165		170		175	
20	Ser	Ile	Asn	Ala	His	Leu	Leu
		180		185		190	
	Asn	Thr	Arg	Thr	Ala	Gln	Gly
		195		200		205	
	Phe	Tyr	Ala	Ile	Asn	Gln	Lys
25		210		215		220	
	Lys	Glu	Leu	Pro	Leu	Ala	Ser
	225		230		235		240
	Glu	Gly	Leu	Leu	Asn	Gln	Leu
		245		250		255	
30	Leu	Asn	Pro	Asn	Ser	Val	His
		260		265		270	
	Ala	Lys	Ile	Ser	Ser	Leu	Pro
		275		280		285	
	Lys	Leu	Asn	His	Phe	Pro	Leu
35		290		295		300	
	Pro	Ile	Glu	Ser	Asn	Ala	Tyr
	305		310		315		320
	Glu	Val	Phe	Leu	Lys	Tyr	Gly
		325		330		335	
40	Phe	Tyr	Ile	Glu	Phe	Asp	Pro
		340		345		350	
	Asp	Asn	Lys	Asp	Gln	Met	Val
		355		360		365	
	Glu	Ser	Phe	Tyr	Ile	His	Ala
45		370		375		380	
	Gly	Phe	Ser	Val	Ser	Asn	Glu
	385		390		395		400
	Ile	Lys	Phe	Lys	Asp	Phe	Gln
		405		410		415	
50	Arg	Ile	Tyr	Arg	Ile	Glu	Phe

420
Ile Leu Val Lys Phe Val
435

425

430

(2) INFORMATION FOR SEQ ID NO:31:

5 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1280 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

10 (ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 66...1223
- (D) OTHER INFORMATION:

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

```
ATCAATACCC CTAAATAAAA AGATATAATG CTGTATTATA AGCTAGTTTT AATTACAATT      60
TTCAA ATG TTA AGG AAA AAC ATT TTA GCT TAC TAT GGG GCG AAT TTT CTC      110
Met Leu Arg Lys Asn Ile Leu Ala Tyr Tyr Gly Ala Asn Phe Leu
      1              5              10              15

20 TTA ATC ATC GCT CAA AGC TTA CCC CAT GCG ATT TTA ACC CCC TTG TTG      158
Leu Ile Ile Ala Gln Ser Leu Pro His Ala Ile Leu Thr Pro Leu Leu
      20              25              30

CTT TCT AAA GGG CTT AGT TTG AGT GAA ATC TTG CTC GTG CAA ACC TTT      206
Leu Ser Lys Gly Leu Ser Leu Ser Glu Ile Leu Leu Val Gln Thr Phe
25      35              40              45

TTT AGC TTT TGC GTG CTA GTG GCT GAA TAC CCA AGC GGC GTT TTA GCG      254
Phe Ser Phe Cys Val Leu Val Ala Glu Tyr Pro Ser Gly Val Leu Ala
      50              55              60

GAT TTG ATG AGC CGA AAA AAT TTA TTC CTG GTT TCT AAT GCC TTT TTA      302
30 Asp Leu Met Ser Arg Lys Asn Leu Phe Leu Val Ser Asn Ala Phe Leu
      65              70              75

ATC GCT AGT TTT TCG TTT GTG CTG TTT TTT GAT AGC TTT ATT TTC ATG      350
Ile Ala Ser Phe Ser Phe Val Leu Phe Phe Asp Ser Phe Ile Phe Met
      80              85              90              95

35 CTT TTA GCG TGG GGG TTG TAT GGT TTG TAT AGC GCA TGC TCT AGC GGC      398
Leu Leu Ala Trp Gly Leu Tyr Gly Leu Tyr Ser Ala Cys Ser Ser Gly
      100              105              110
```

	ACG ATT GAA GCT TCA CTC ATC ACA GAC ATT AAG GAA AAC AAA AAA GAT	446
	Thr Ile Glu Ala Ser Leu Ile Thr Asp Ile Lys Glu Asn Lys Lys Asp	
	115 120 125	
	TTA TCC AAG TTT TTA GCC AAA AAC AAT CAA ATT ACT TAT TTA GGC ATG	494
5	Leu Ser Lys Phe Leu Ala Lys Asn Asn Gln Ile Thr Tyr Leu Gly Met	
	130 135 140	
	ATT ATA GGG AGT TCT TTG GGA TCG TTT TTG TAT CTC AAA GTC CAT GCG	542
	Ile Ile Gly Ser Ser Leu Gly Ser Phe Leu Tyr Leu Lys Val His Ala	
	145 150 155	
10	ATG CTG TAT ATT GTG GGG ATT TTT TTA ATC ATG CTC TGT GTG CTA ACG	590
	Met Leu Tyr Ile Val Gly Ile Phe Leu Ile Met Leu Cys Val Leu Thr	
	160 165 170 175	
	ATC ATT TTT TAT TTT AAA GAG AAA GAA GGG GAT TTT AAA AGC CAA AAA	638
15	Ile Ile Phe Tyr Phe Lys Glu Lys Glu Gly Asp Phe Lys Ser Gln Lys	
	180 185 190	
	AGC CTG AAA CTC CTT AAA GAG CAA GTC AAA GGC AGT CTT AAA GAG CTT	686
	Ser Leu Lys Leu Leu Lys Glu Gln Val Lys Gly Ser Leu Lys Glu Leu	
	195 200 205	
	AAA GAT AAC CCC AAA CTT AAA ATT CTG TTA GTG GGG CAT TTG ATT ACG	734
20	Lys Asp Asn Pro Lys Leu Lys Ile Leu Leu Val Gly His Leu Ile Thr	
	210 215 220	
	CCC GTC TTT TTT ATG AGC CAT TTT CAA ATG TGG CAA GCG TAT TTT TTA	782
	Pro Val Phe Phe Met Ser His Phe Gln Met Trp Gln Ala Tyr Phe Leu	
	225 230 235	
25	AAA CAA GGC GTT AAA GAG CAA TAC CTT TTT GTG TTT TAT ATC GCT TTT	830
	Lys Gln Gly Val Lys Glu Gln Tyr Leu Phe Val Phe Tyr Ile Ala Phe	
	240 245 250 255	
	CAA GTG ATT TCT ATT CTC ATT CAT TTT TTA AAA GCC TCT AGT TAT AGC	878
30	Gln Val Ile Ser Ile Leu Ile His Phe Leu Lys Ala Ser Ser Tyr Ser	
	260 265 270	
	CAA AAA ATC GCC TTG AGT TCG CTT GTG GTG TTG TTA GGC GTT AGC CCC	926
	Gln Lys Ile Ala Leu Ser Ser Leu Val Val Leu Leu Gly Val Ser Pro	
	275 280 285	
	TTA TTG CTT AGC AAT ATC CCT TAT TGT TTC ATA GGG GTG TAT GCG CTC	974
35	Leu Leu Leu Ser Asn Ile Pro Tyr Cys Phe Ile Gly Val Tyr Ala Leu	
	290 295 300	
	ATG GTG GCG TTT TTC ACT TAC ATG AGC TAT TGC TTA AAC TAT CAA TTC	1022
	Met Val Ala Phe Phe Thr Tyr Met Ser Tyr Cys Leu Asn Tyr Gln Phe	

305

310

315

TCC AAA TTC GTT TCT AAA AAC AAC ATT TCC TCG CTC TCA TCG CTT TTA 1070
 Ser Lys Phe Val Ser Lys Asn Asn Ile Ser Ser Leu Ser Ser Leu Leu
 320 325 330 335

5 TCA AGC TGT GTG CGC GTG GTC TCT GTG CTA ATC TTA TCG CTC AGC AGT 1118
 Ser Ser Cys Val Arg Val Val Ser Val Leu Ile Leu Ser Leu Ser Ser
 340 345 350

10 CTG GAA CTG CGT TAC TTC TCA CCC CTA ACT ATC ATA ACC ATG CAT TTT 1166
 Leu Glu Leu Arg Tyr Phe Ser Pro Leu Thr Ile Ile Thr Met His Phe
 355 360 365

GCC TTG ACG CTT ATC ATC CTC TTT TTC TTT TTG TAT AAG GCT AAG CCG 1214
 Ala Leu Thr Leu Ile Ile Leu Phe Phe Phe Leu Tyr Lys Ala Lys Pro
 370 375 380

15 TTT GAT GAG TGAGCGGCTT TAAGAGTGCA ACCTTTTAGC GATTTCCTATA GCAACATCA 1272
 Phe Asp Glu
 385

TAGCCATG 1280

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

20 (A) LENGTH: 386 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

25 (v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

Met Leu Arg Lys Asn Ile Leu Ala Tyr Tyr Gly Ala Asn Phe Leu Leu
 1 5 10 15
 Ile Ile Ala Gln Ser Leu Pro His Ala Ile Leu Thr Pro Leu Leu Leu
 30 20 25 30
 Ser Lys Gly Leu Ser Leu Ser Glu Ile Leu Leu Val Gln Thr Phe Phe
 35 40 45
 Ser Phe Cys Val Leu Val Ala Glu Tyr Pro Ser Gly Val Leu Ala Asp
 50 55 60
 35 Leu Met Ser Arg Lys Asn Leu Phe Leu Val Ser Asn Ala Phe Leu Ile
 65 70 75 80
 Ala Ser Phe Ser Phe Val Leu Phe Phe Asp Ser Phe Ile Phe Met Leu
 85 90 95
 Leu Ala Trp Gly Leu Tyr Gly Leu Tyr Ser Ala Cys Ser Ser Gly Thr

(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

5	ATTAAATATG ACTATATACA CTACAACAAT AAGATTTTGA AAGGTTGGTA	ATG GAA	56
		Met Glu	
		1	
	TCA GTA AAA ACA GGA AAA ACA AAT AAG GTT GGC AAG AAT ACA GAG ATG		104
	Ser Val Lys Thr Gly Lys Thr Asn Lys Val Gly Lys Asn Thr Glu Met		
	5 10 15		
10	GCT AAT ACA AAG GCA AAT AAA GAG GCT CAT TTT AAA CAA GCG AGC ACC		152
	Ala Asn Thr Lys Ala Asn Lys Glu Ala His Phe Lys Gln Ala Ser Thr		
	20 25 30		
	ATT ACA AAT ATA ATC AGA TCA ATT CGT GGG ATT TTT ACA AAA ATT GCA		200
	Ile Thr Asn Ile Ile Arg Ser Ile Arg Gly Ile Phe Thr Lys Ile Ala		
15	35 40 45 50		
	AAG AAA GTT AGA GGA CTT GTA AAA AAA CAC CCC AAG AAA AGC AGT GCG		248
	Lys Lys Val Arg Gly Leu Val Lys Lys His Pro Lys Lys Ser Ser Ala		
	55 60 65		
	GCA TTA GTA GTA TTG ACC CAT ATT GCG TGC AAG AAA GCG AAA GAA TTA		296
20	Ala Leu Val Val Leu Thr His Ile Ala Cys Lys Lys Ala Lys Glu Leu		
	70 75 80		
	GAC GAT AAA GTC CAA GAT AAA TCC AAA CAA GCT GAA AAA GAA AAT CAA		344
	Asp Asp Lys Val Gln Asp Lys Ser Lys Gln Ala Glu Lys Glu Asn Gln		
	85 90 95		
25	ATC AAT TGG TGG AAA TAT TCA GGA TTA ACA ATA GCG ACA AGT TTA TTA		392
	Ile Asn Trp Trp Lys Tyr Ser Gly Leu Thr Ile Ala Thr Ser Leu Leu		
	100 105 110		
	TTA GCC GCT TGT AGC ACT GGT GAT GTT AGT GAA CAA ATA GAA CTA GAA		440
	Leu Ala Ala Cys Ser Thr Gly Asp Val Ser Glu Gln Ile Glu Leu Glu		
30	115 120 125 130		
	CAA GAA AAA CAA AAG ACG AGC AAT ATA GAG ACT AAC AAT CAA ATA AAA		488
	Gln Glu Lys Gln Lys Thr Ser Asn Ile Glu Thr Asn Asn Gln Ile Lys		
	135 140 145		
	GTA GAA CAA GAA AAA CAA AAG ACA AGC AAT ATA GAG ACT AAT AAT CAA		536
35	Val Glu Gln Glu Lys Gln Lys Thr Ser Asn Ile Glu Thr Asn Asn Gln		
	150 155 160		
	ATA AAA GTA GAA CAA GAA CAA CAG AAA ACA GAA CAA GAA MGA CAG AAA		584

Ile Lys Val Glu Gln Glu Gln Gln Lys Thr Glu Gln Glu Xaa Gln Lys
165 170 175

	ACA GAA CAA GAA AGA CAG AAG ACA GAA CAA GAA AAA CAA AAG ACC ATT	632
5	Thr Glu Gln Glu Arg Gln Lys Thr Glu Gln Glu Lys Gln Lys Thr Ile 180 185 190	
	AAA ACA CAG AAA GAT TTC ATT AAA TAT GTA GAA CAA AAT TGC CAA GAA	680
	Lys Thr Gln Lys Asp Phe Ile Lys Tyr Val Glu Gln Asn Cys Gln Glu 195 200 205 210	
10	AAT CAT AAT CAA TTC TTT ATT GAA AAA GGA GGA ATT AAG GCT GGT ATT	728
	Asn His Asn Gln Phe Phe Ile Glu Lys Gly Gly Ile Lys Ala Gly Ile 215 220 225	
	GGT ATA GAA GTA GAA GCT GAA TGC AAA ACC CCT AAA CCT GCA AAA ACC	776
	Gly Ile Glu Val Glu Ala Glu Cys Lys Thr Pro Lys Pro Ala Lys Thr 230 235 240	
15	AAT CAA ACC CCT ATC CAG CCA AAA CAC CTC CCA AAC TCT AAA CAA CCC	824
	Asn Gln Thr Pro Ile Gln Pro Lys His Leu Pro Asn Ser Lys Gln Pro 245 250 255	
20	CGC TCT CAA AGA GGA TCA AAA GCG CAA GAG CTT ATC GCT TAT TTG CAA	872
	Arg Ser Gln Arg Gly Ser Lys Ala Gln Glu Leu Ile Ala Tyr Leu Gln 260 265 270	
	AAA GAG CTA GAA TCT CTG CCC TAT TCA CAA AAA GCT ATC GCT AAA CAA	920
	Lys Glu Leu Glu Ser Leu Pro Tyr Ser Gln Lys Ala Ile Ala Lys Gln 275 280 285 290	
25	GTG GAT TTT TAT AGA CCA AGT TCT ATC GCT TAT TTA GAA CTA GAC CCT	968
	Val Asp Phe Tyr Arg Pro Ser Ser Ile Ala Tyr Leu Glu Leu Asp Pro 295 300 305	
	AGA GAT TTT AAT GTT ACA GAA GAA TGG CAA AAA GAA AAT TTA AAA ATA	1016
	Arg Asp Phe Asn Val Thr Glu Glu Trp Gln Lys Glu Asn Leu Lys Ile 310 315 320	
30	CGC TCT AAA GCT CAA GCT AAA ATG CTT GAA ATG AGG AGT TTA AAA CCA	1064
	Arg Ser Lys Ala Gln Ala Lys Met Leu Glu Met Arg Ser Leu Lys Pro 325 330 335	
35	GAC TCA CAA GCC CAC CTT TCA ACC TCT CAA AGC CTT TTG TTC GTT CAA	1112
	Asp Ser Gln Ala His Leu Ser Thr Ser Gln Ser Leu Leu Phe Val Gln 340 345 350	
	AAA ATA TTT GCT GAT GTT AAT AAA GAA ATA AAA GTA GTT GCT AAT ACT	1160
	Lys Ile Phe Ala Asp Val Asn Lys Glu Ile Lys Val Val Ala Asn Thr 355 360 365 370	

GAA AAG AAA GCA GAA AAA GCG GGT TAT GGT TAT AGT AAA AGG ATG TAGGC 1210
 Glu Lys Lys Ala Glu Lys Ala Gly Tyr Gly Tyr Ser Lys Arg Met
 375 380 385

ATAAGAAAAC ACCATAAAAT CGTTCCTAGC TTATTTATAG TATTTTAAAA ACTC 1264

5 (2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 385 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

10 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

15 Met Glu Ser Val Lys Thr Gly Lys Thr Asn Lys Val Gly Lys Asn Thr
 1 5 10 15
 Glu Met Ala Asn Thr Lys Ala Asn Lys Glu Ala His Phe Lys Gln Ala
 20 25 30
 Ser Thr Ile Thr Asn Ile Ile Arg Ser Ile Arg Gly Ile Phe Thr Lys
 35 40 45
 20 Ile Ala Lys Lys Val Arg Gly Leu Val Lys Lys His Pro Lys Lys Ser
 50 55 60
 Ser Ala Ala Leu Val Val Leu Thr His Ile Ala Cys Lys Lys Ala Lys
 65 70 75 80
 Glu Leu Asp Asp Lys Val Gln Asp Lys Ser Lys Gln Ala Glu Lys Glu
 25 85 90 95
 Asn Gln Ile Asn Trp Trp Lys Tyr Ser Gly Leu Thr Ile Ala Thr Ser
 100 105 110
 Leu Leu Leu Ala Ala Cys Ser Thr Gly Asp Val Ser Glu Gln Ile Glu
 115 120 125
 30 Leu Glu Gln Glu Lys Gln Lys Thr Ser Asn Ile Glu Thr Asn Asn Gln
 130 135 140
 Ile Lys Val Glu Gln Glu Lys Gln Lys Thr Ser Asn Ile Glu Thr Asn
 145 150 155 160
 Asn Gln Ile Lys Val Glu Gln Glu Gln Lys Thr Glu Gln Glu Xaa
 35 165 170 175
 Gln Lys Thr Glu Gln Glu Arg Gln Lys Thr Glu Gln Glu Lys Gln Lys
 180 185 190
 Thr Ile Lys Thr Gln Lys Asp Phe Ile Lys Tyr Val Glu Gln Asn Cys
 195 200 205
 40 Gln Glu Asn His Asn Gln Phe Phe Ile Glu Lys Gly Gly Ile Lys Ala
 210 215 220
 Gly Ile Gly Ile Glu Val Glu Ala Glu Cys Lys Thr Pro Lys Pro Ala
 225 230 235 240
 Lys Thr Asn Gln Thr Pro Ile Gln Pro Lys His Leu Pro Asn Ser Lys

245 250 255
 Gln Pro Arg Ser Gln Arg Gly Ser Lys Ala Gln Glu Leu Ile Ala Tyr
 260 265 270
 5 Leu Gln Lys Glu Leu Glu Ser Leu Pro Tyr Ser Gln Lys Ala Ile Ala
 275 280 285
 Lys Gln Val Asp Phe Tyr Arg Pro Ser Ser Ile Ala Tyr Leu Glu Leu
 290 295 300
 Asp Pro Arg Asp Phe Asn Val Thr Glu Glu Trp Gln Lys Glu Asn Leu
 305 310 315 320
 10 Lys Ile Arg Ser Lys Ala Gln Ala Lys Met Leu Glu Met Arg Ser Leu
 325 330 335
 Lys Pro Asp Ser Gln Ala His Leu Ser Thr Ser Gln Ser Leu Leu Phe
 340 345 350
 Val Gln Lys Ile Phe Ala Asp Val Asn Lys Glu Ile Lys Val Val Ala
 15 355 360 365
 Asn Thr Glu Lys Lys Ala Glu Lys Ala Gly Tyr Gly Tyr Ser Lys Arg
 370 375 380
 Met
 385

20 (2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 410 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 25 (D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence
 (B) LOCATION: 62...340
 (D) OTHER INFORMATION:

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

ATTCATTTAC TTTTGAGAAA TATAATTCTC TCGCTTTTAA GATCATCACA AGGAGTTTCG 60
 T ATG AAA AAG CAA ATC TTG ACA GGT GTT TTA TTA TCA GTT TTG GCA GTG 109
 Met Lys Lys Gln Ile Leu Thr Gly Val Leu Leu Ser Val Leu Ala Val
 35 1 5 10 15
 AGT TCT GCA TAC GCT CAC AAA GAT AAA AAA GAC GCC AAA AAA CCT AAA 157
 Ser Ser Ala Tyr Ala His Lys Asp Lys Lys Asp Ala Lys Lys Pro Lys
 20 25 30
 TTT AGC ACA GAA TTA GTC GTG GCT CAA AAC GAC AAA AAA GAC GCT AAA 205
 40 Phe Ser Thr Glu Leu Val Val Ala Gln Asn Asp Lys Lys Asp Ala Lys
 35 40 45

AAA CCT AAA TTT AGC ACA GAA TTA GTC GTG GCT CAA AAC GAC AAA AAA 253
 Lys Pro Lys Phe Ser Thr Glu Leu Val Val Ala Gln Asn Asp Lys Lys
 50 55 60

GAC GCT AAA AAA CCT AAA TTT AGC ACA GAA TTA GTC GTG GCT CAA AAC 301
 5 Asp Ala Lys Lys Pro Lys Phe Ser Thr Glu Leu Val Val Ala Gln Asn
 65 70 75 80

GAC AAA AAA GAC GCT AAA AAA CCT AAA AAC TCA GTG GTC TAATGGCTTT GA 352
 Asp Lys Lys Asp Ala Lys Lys Pro Lys Asn Ser Val Val
 85 90

10 CTCTAAAAAA GCGTTTTTAA AAACGCTTTT TTGGATATTA TCCTATAATT TCCTACCA 410

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 93 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

20 Met Lys Lys Gln Ile Leu Thr Gly Val Leu Leu Ser Val Leu Ala Val
 1 5 10 15
 Ser Ser Ala Tyr Ala His Lys Asp Lys Lys Asp Ala Lys Lys Pro Lys
 20 25 30
 Phe Ser Thr Glu Leu Val Val Ala Gln Asn Asp Lys Lys Asp Ala Lys
 25 35 40 45
 Lys Pro Lys Phe Ser Thr Glu Leu Val Val Ala Gln Asn Asp Lys Lys
 50 55 60
 Asp Ala Lys Lys Pro Lys Phe Ser Thr Glu Leu Val Val Ala Gln Asn
 65 70 75 80
 30 Asp Lys Lys Asp Ala Lys Lys Pro Lys Asn Ser Val Val
 85 90

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2097 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence
 (B) LOCATION: 67...2046
 (D) OTHER INFORMATION:

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

	TAAAAACCCC TATCATAGGG CGTGGCATGA AGAAAAAGC AAAAGTCTTT TGGTATTGTT	60
	TTAATC ATG ATT TAT TGG TTG TAT TTG GCG GTC TTT TTT TTG TTG AGC	108
	Met Ile Tyr Trp Leu Tyr Leu Ala Val Phe Phe Leu Leu Ser	
	1 5 10	
10	GCA TTA GAC GCT AAA GAA ATC GCT ATG CAA CGA TTT GAC AAA CAA AAC	156
	Ala Leu Asp Ala Lys Glu Ile Ala Met Gln Arg Phe Asp Lys Gln Asn	
	15 20 25 30	
	CAT AAG ATT TTT GAA ATC CTT GCG GAT AAA GTG AGC GCT AAA GAC AAT	204
	His Lys Ile Phe Glu Ile Leu Ala Asp Lys Val Ser Ala Lys Asp Asn	
15	35 40 45	
	GTG ATA ACC GCA TCA GGG AAT GCG ATC TTA TTG AAT TAT GAT GTG TAT	252
	Val Ile Thr Ala Ser Gly Asn Ala Ile Leu Leu Asn Tyr Asp Val Tyr	
	50 55 60	
	ATT CTA GCG GAC AAG GTG CGT TAT GAC ACT AAA ACC AAA GAA GCG TTA	300
20	Ile Leu Ala Asp Lys Val Arg Tyr Asp Thr Lys Thr Lys Glu Ala Leu	
	65 70 75	
	TTA GAG GGG AAT ATC AAG GTT TAT AGG GGC GAG GGT TTG CTC GTT AAA	348
	Leu Glu Gly Asn Ile Lys Val Tyr Arg Gly Glu Gly Leu Leu Val Lys	
	80 85 90	
25	ACC GAT TAC GTG AAA TTG AGT TTG AAT GAA AAA TAT GAA ATC ATT TTC	396
	Thr Asp Tyr Val Lys Leu Ser Leu Asn Glu Lys Tyr Glu Ile Ile Phe	
	95 100 105 110	
	CCC TTT TAT GTC CAA GAC AGC GTG AGC GGG ATT TGG GTG AGC GCG GAT	444
	Pro Phe Tyr Val Gln Asp Ser Val Ser Gly Ile Trp Val Ser Ala Asp	
30	115 120 125	
	ATT GCC AGC GGA AAG GAT CAA AAA TAT AAG GTT AAA AAC ATG AGC ACT	492
	Ile Ala Ser Gly Lys Asp Gln Lys Tyr Lys Val Lys Asn Met Ser Thr	
	130 135 140	
	TCA GGG TGC AGC ATT GAT AAC CCC ATT TGG CAT GTC AAT GCG ACT TCA	540
35	Ser Gly Cys Ser Ile Asp Asn Pro Ile Trp His Val Asn Ala Thr Ser	
	145 150 155	
	GGC TCA TTC AAC ATG CAA AAA TCG CAT TTG TCT ATG TGG AAT CCT AAG	588
	Gly Ser Phe Asn Met Gln Lys Ser His Leu Ser Met Trp Asn Pro Lys	

160

165

170

ATC TAT GTC GGT GAT ATT CCT GTA TTG TAT TTG CCC TAT ATT TTC ATG 636
 Ile Tyr Val Gly Asp Ile Pro Val Leu Tyr Leu Pro Tyr Ile Phe Met
 175 180 185 190

5 TCC ACG AGC AAT AAA AGA ACT ACT GGG TTT TTA TAC CCT GAG TTT GGC 684
 Ser Thr Ser Asn Lys Arg Thr Thr Gly Phe Leu Tyr Pro Glu Phe Gly
 195 200 205

10 ACT TCC AAC TTA GAC GGC TTT ATT TAT TTG CAA CCC TTT TAT TTA GCC 732
 Thr Ser Asn Leu Asp Gly Phe Ile Tyr Leu Gln Pro Phe Tyr Leu Ala
 210 215 220

CCC AAA AAC TCA TGG GAT ATG ACC TTT ACC CCA CAA ATC CGC TAT AAA 780
 Pro Lys Asn Ser Trp Asp Met Thr Phe Thr Pro Gln Ile Arg Tyr Lys
 225 230 235

15 AGG GGT TTT GGC TTG AAT TTT GAA GCG CGC TAC ATT AAC TCT AAA AAC 828
 Arg Gly Phe Gly Leu Asn Phe Glu Ala Arg Tyr Ile Asn Ser Lys Asn
 240 245 250

GAC AGG TTT TTA TTC AAC GCG CGC TAT TTT AGG AAT TAC ACC CAA TAT 876
 Asp Arg Phe Leu Phe Asn Ala Arg Tyr Phe Arg Asn Tyr Thr Gln Tyr
 255 260 265 270

20 GTC AAA CGC TAC GAT TTG AGG AAT CAA AAT ATC TAC GGG TTT GAA TTT 924
 Val Lys Arg Tyr Asp Leu Arg Asn Gln Asn Ile Tyr Gly Phe Glu Phe
 275 280 285

25 TTA AGC TCT AGC AGG GAC ACT TTA CAA AAA TAC TTC CAC CTT AAG TCT 972
 Leu Ser Ser Ser Arg Asp Thr Leu Gln Lys Tyr Phe His Leu Lys Ser
 290 295 300

AAT ATT GAC AAC GGG CAT TAC ATT GAC TTT TTA TAC ATG AAC GAT TTG 1020
 Asn Ile Asp Asn Gly His Tyr Ile Asp Phe Leu Tyr Met Asn Asp Leu
 305 310 315

30 GAC TAT GTG CGT TTT GAA AAG GTT AAT AAG CGT ATC ACA GAC GCC ACG 1068
 Asp Tyr Val Arg Phe Glu Lys Val Asn Lys Arg Ile Thr Asp Ala Thr
 320 325 330

CAC ATG TCT AGG GCG AAT TAC TAT TTG CAA ACA GAA AAC AAT TAT TAC 1116
 His Met Ser Arg Ala Asn Tyr Tyr Leu Gln Thr Glu Asn Asn Tyr Tyr
 335 340 345 350

35 GGC TTG AAT ATC AAG TAT TTT TTA AAC CTG AAT AAA ATC AAC AAT AAC 1164
 Gly Leu Asn Ile Lys Tyr Phe Leu Asn Leu Asn Lys Ile Asn Asn Asn
 355 360 365

	CGC ACT TTC CAA TCT GTC CCT AAT TTG CAA TAC CAT AAA TAT TTA AAT	1212
	Arg Thr Phe Gln Ser Val Pro Asn Leu Gln Tyr His Lys Tyr Leu Asn	
	370 375 380	
5	TCT TTG TAT TTT AGA AAT TTG TTG TAT TCG GTG GAT TAT CAG TTT AGA	1260
	Ser Leu Tyr Phe Arg Asn Leu Leu Tyr Ser Val Asp Tyr Gln Phe Arg	
	385 390 395	
	AAC ACC GCA AGA GAG ATT GGT TAT GGC TAT GTG CAA AAC GCT TTG AAT	1308
	Asn Thr Ala Arg Glu Ile Gly Tyr Gly Tyr Val Gln Asn Ala Leu Asn	
	400 405 410	
10	GTG CCG GTG GGC TTG CAA TTT TCT TTG TTT AAA AAG TAT TTG TCT TTA	1356
	Val Pro Val Gly Leu Gln Phe Ser Leu Phe Lys Lys Tyr Leu Ser Leu	
	415 420 425 430	
15	GGG CTT TGG AAT GAT CTC CAA CTA TCT AAT GTG GCT TTA ATG CAA TCT	1404
	Gly Leu Trp Asn Asp Leu Gln Leu Ser Asn Val Ala Leu Met Gln Ser	
	435 440 445	
	AAA AAT TCC TTC GTG CCT ACG ATC CCT AAT GAA TCA AGG GAA TTT GGG	1452
	Lys Asn Ser Phe Val Pro Thr Ile Pro Asn Glu Ser Arg Glu Phe Gly	
	450 455 460	
20	AAT TTT GTG TCT TCA AAT TTT TCC ATG TAT GTC AAT ACG GAT TTG GCT	1500
	Asn Phe Val Ser Ser Asn Phe Ser Met Tyr Val Asn Thr Asp Leu Ala	
	465 470 475	
	AGA GAA TAC AAC AAG CTT TTC CAC ACG ATC CAA CTA GAA GCG ATT TTC	1548
	Arg Glu Tyr Asn Lys Leu Phe His Thr Ile Gln Leu Glu Ala Ile Phe	
	480 485 490	
25	AAC ATC CCT TAT TAC ACC TTT AAA AAC GGC TTA TTT TCT CAA AAC ATG	1596
	Asn Ile Pro Tyr Tyr Thr Phe Lys Asn Gly Leu Phe Ser Gln Asn Met	
	495 500 505 510	
30	TAT GCT TTA AGC GCG CAA GCC TTA AAC AGC TAC ACT TCG CCT TTA TTG	1644
	Tyr Ala Leu Ser Ala Gln Ala Leu Asn Ser Tyr Thr Ser Pro Leu Leu	
	515 520 525	
	AGA GAT TAT GAT TAT CAA GGG CGT TTG TAT GAC TCG GTG TGG AAT CCT	1692
	Arg Asp Tyr Asp Tyr Gln Gly Arg Leu Tyr Asp Ser Val Trp Asn Pro	
	530 535 540	
35	AGC AGT ATT TTA CCT AGC AAT GCG AGC AAC AAG ACG GTG GAT TTA ACC	1740
	Ser Ser Ile Leu Pro Ser Asn Ala Ser Asn Lys Thr Val Asp Leu Thr	
	545 550 555	
	CTA ACG CAA TAC CTT TAT GGC TTA GGG GGG CAA GAG TTA TTG TAT TTT	1788
	Leu Thr Gln Tyr Leu Tyr Gly Leu Gly Gly Gln Glu Leu Leu Tyr Phe	

560

565

570

AAA ATA TCG CAA CTC ATC AAT CTT GAC GAT AAA GTT TCG CCC TTT AGA 1836
 Lys Ile Ser Gln Leu Ile Asn Leu Asp Asp Lys Val Ser Pro Phe Arg
 575 580 585 590

5 ATG CCA CTA GAG AGC AAG ATC GGG TTT TCG CCC TTA ACG GGA TTG AAC 1884
 Met Pro Leu Glu Ser Lys Ile Gly Phe Ser Pro Leu Thr Gly Leu Asn
 595 600 605

10 ATC TTT GGG AAT GTC TTT TAT TCG TTT TAT CAA AAC CGC TTA GAA GAA 1932
 Ile Phe Gly Asn Val Phe Tyr Ser Phe Tyr Gln Asn Arg Leu Glu Glu
 610 615 620

ATC TCT GTG AAC GCC AAT TAC CAA CGC AAG TTT TTA AGC TTT AAC CTC 1980
 Ile Ser Val Asn Ala Asn Tyr Gln Arg Lys Phe Leu Ser Phe Asn Leu
 625 630 635

15 TCT TAT TTT TTA AAA AAC AAT TTT AGC AGT GGG ATT AAT AGC ATT GTA 2028
 Ser Tyr Phe Leu Lys Asn Asn Phe Ser Ser Gly Ile Asn Ser Ile Val
 640 645 650

GAA AAT CTG CGG ATT ATT TAAAGGCGGG TTTTAGCAAC GACTTTGGCT ATTTTTC 2084
 Glu Asn Leu Arg Ile Ile
 655 660

20 ATGAGCGCGG ATG 2097

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 660 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

30 Met Ile Tyr Trp Leu Tyr Leu Ala Val Phe Phe Leu Leu Ser Ala Leu
 1 5 10 15
 Asp Ala Lys Glu Ile Ala Met Gln Arg Phe Asp Lys Gln Asn His Lys
 20 25 30
 Ile Phe Glu Ile Leu Ala Asp Lys Val Ser Ala Lys Asp Asn Val Ile
 35 40 45
 Thr Ala Ser Gly Asn Ala Ile Leu Leu Asn Tyr Asp Val Tyr Ile Leu
 50 55 60
 Ala Asp Lys Val Arg Tyr Asp Thr Lys Thr Lys Glu Ala Leu Leu Glu

	65				70					75				80
	Gly	Asn	Ile	Lys	Val	Tyr	Arg	Gly	Glu	Gly	Leu	Leu	Val	Lys Thr Asp
					85					90				95
	Tyr	Val	Lys	Leu	Ser	Leu	Asn	Glu	Lys	Tyr	Glu	Ile	Ile	Phe Pro Phe
5				100					105					110
	Tyr	Val	Gln	Asp	Ser	Val	Ser	Gly	Ile	Trp	Val	Ser	Ala	Asp Ile Ala
			115					120					125	
	Ser	Gly	Lys	Asp	Gln	Lys	Tyr	Lys	Val	Lys	Asn	Met	Ser	Thr Ser Gly
		130					135					140		
10	Cys	Ser	Ile	Asp	Asn	Pro	Ile	Trp	His	Val	Asn	Ala	Thr	Ser Gly Ser
						150					155			160
	Phe	Asn	Met	Gln	Lys	Ser	His	Leu	Ser	Met	Trp	Asn	Pro	Lys Ile Tyr
				165						170				175
	Val	Gly	Asp	Ile	Pro	Val	Leu	Tyr	Leu	Pro	Tyr	Ile	Phe	Met Ser Thr
15				180					185				190	
	Ser	Asn	Lys	Arg	Thr	Thr	Gly	Phe	Leu	Tyr	Pro	Glu	Phe	Gly Thr Ser
			195				200					205		
	Asn	Leu	Asp	Gly	Phe	Ile	Tyr	Leu	Gln	Pro	Phe	Tyr	Leu	Ala Pro Lys
		210					215					220		
20	Asn	Ser	Trp	Asp	Met	Thr	Phe	Thr	Pro	Gln	Ile	Arg	Tyr	Lys Arg Gly
							230				235			240
	Phe	Gly	Leu	Asn	Phe	Glu	Ala	Arg	Tyr	Ile	Asn	Ser	Lys	Asn Asp Arg
				245					250					255
	Phe	Leu	Phe	Asn	Ala	Arg	Tyr	Phe	Arg	Asn	Tyr	Thr	Gln	Tyr Val Lys
25				260					265				270	
	Arg	Tyr	Asp	Leu	Arg	Asn	Gln	Asn	Ile	Tyr	Gly	Phe	Glu	Phe Leu Ser
			275				280					285		
	Ser	Ser	Arg	Asp	Thr	Leu	Gln	Lys	Tyr	Phe	His	Leu	Lys	Ser Asn Ile
		290					295					300		
30	Asp	Asn	Gly	His	Tyr	Ile	Asp	Phe	Leu	Tyr	Met	Asn	Asp	Leu Asp Tyr
							310				315			320
	Val	Arg	Phe	Glu	Lys	Val	Asn	Lys	Arg	Ile	Thr	Asp	Ala	Thr His Met
				325					330					335
	Ser	Arg	Ala	Asn	Tyr	Tyr	Leu	Gln	Thr	Glu	Asn	Asn	Tyr	Tyr Gly Leu
35				340					345				350	
	Asn	Ile	Lys	Tyr	Phe	Leu	Asn	Leu	Asn	Lys	Ile	Asn	Asn	Asn Arg Thr
		355					360					365		
	Phe	Gln	Ser	Val	Pro	Asn	Leu	Gln	Tyr	His	Lys	Tyr	Leu	Asn Ser Leu
		370					375				380			
40	Tyr	Phe	Arg	Asn	Leu	Leu	Tyr	Ser	Val	Asp	Tyr	Gln	Phe	Arg Asn Thr
						390					395			400
	Ala	Arg	Glu	Ile	Gly	Tyr	Gly	Tyr	Val	Gln	Asn	Ala	Leu	Asn Val Pro
				405					410					415
	Val	Gly	Leu	Gln	Phe	Ser	Leu	Phe	Lys	Lys	Tyr	Leu	Ser	Leu Gly Leu
45				420					425				430	
	Trp	Asn	Asp	Leu	Gln	Leu	Ser	Asn	Val	Ala	Leu	Met	Gln	Ser Lys Asn
		435					440					445		
	Ser	Phe	Val	Pro	Thr	Ile	Pro	Asn	Glu	Ser	Arg	Glu	Phe	Gly Asn Phe
		450					455				460			
50	Val	Ser	Ser	Asn	Phe	Ser	Met	Tyr	Val	Asn	Thr	Asp	Leu	Ala Arg Glu

	465		470		475		480									
	Tyr	Asn	Lys	Leu	Phe	His	Thr	Ile	Gln	Leu	Glu	Ala	Ile	Phe	Asn	Ile
					485				490						495	
5	Pro	Tyr	Tyr	Thr	Phe	Lys	Asn	Gly	Leu	Phe	Ser	Gln	Asn	Met	Tyr	Ala
			500						505					510		
	Leu	Ser	Ala	Gln	Ala	Leu	Asn	Ser	Tyr	Thr	Ser	Pro	Leu	Leu	Arg	Asp
			515					520					525			
	Tyr	Asp	Tyr	Gln	Gly	Arg	Leu	Tyr	Asp	Ser	Val	Trp	Asn	Pro	Ser	Ser
		530					535					540				
10	Ile	Leu	Pro	Ser	Asn	Ala	Ser	Asn	Lys	Thr	Val	Asp	Leu	Thr	Leu	Thr
		545				550					555					560
	Gln	Tyr	Leu	Tyr	Gly	Leu	Gly	Gly	Gln	Glu	Leu	Leu	Tyr	Phe	Lys	Ile
					565					570						575
	Ser	Gln	Leu	Ile	Asn	Leu	Asp	Asp	Lys	Val	Ser	Pro	Phe	Arg	Met	Pro
15				580					585					590		
	Leu	Glu	Ser	Lys	Ile	Gly	Phe	Ser	Pro	Leu	Thr	Gly	Leu	Asn	Ile	Phe
			595					600					605			
	Gly	Asn	Val	Phe	Tyr	Ser	Phe	Tyr	Gln	Asn	Arg	Leu	Glu	Glu	Ile	Ser
		610					615					620				
20	Val	Asn	Ala	Asn	Tyr	Gln	Arg	Lys	Phe	Leu	Ser	Phe	Asn	Leu	Ser	Tyr
		625				630					635					640
	Phe	Leu	Lys	Asn	Asn	Phe	Ser	Ser	Gly	Ile	Asn	Ser	Ile	Val	Glu	Asn
				645					650						655	
	Leu	Arg	Ile	Ile												
25				660												

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 961 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 168...764
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

	ATGCCGATTA	AATGCATGCT	GATTAAATGA	ATGAAAAGAG	TCCAAACCAC	CGCCTTTAAC	60
	GCACCACGCT	TGAAATTAAA	ACTAAATTTT	AGTGTATTCT	TAGCAAATTT	TAGATAAGAT	120
40	CAAGCGTGAT	TTTTTCTAAA	TTTGTAGGCAT	TTAAGGAATC	AGTGTTT	ATG ACA AGC	176
						Met Thr Ser	
						1	
	GCT CTG TTA	GGC TTA CAA	ATT GTT TTA	GCG GTA TTG	ATT GTG GTG GTG		224

	Ala	Leu	Leu	Gly	Leu	Gln	Ile	Val	Leu	Ala	Val	Leu	Ile	Val	Val	Val	
	5						10					15					
	GTT	TTG	TTG	CAA	AAA	AGT	TCT	AGC	ATC	GGC	TTA	GGG	GCT	TAT	AGC	GGG	272
	Val	Leu	Leu	Gln	Lys	Ser	Ser	Ser	Ile	Gly	Leu	Gly	Ala	Tyr	Ser	Gly	
5	20					25					30					35	
	AGT	AAT	GAG	TCT	TTA	TTT	GGC	GCT	AAA	GGG	CCT	GCA	AGC	TTT	ATG	GCG	320
	Ser	Asn	Glu	Ser	Leu	Phe	Gly	Ala	Lys	Gly	Pro	Ala	Ser	Phe	Met	Ala	
					40					45					50		
	AAA	TTA	ACC	ATG	TTT	TTA	GGG	CTG	TTA	TTT	GTC	ATC	AAC	ACC	ATC	GCT	368
10	Lys	Leu	Thr	Met	Phe	Leu	Gly	Leu	Leu	Phe	Val	Ile	Asn	Thr	Ile	Ala	
				55					60					65			
	TTG	GGC	TAT	TTT	TAC	AAC	AAA	GAA	TAC	GGC	AAG	AGC	GTT	TTA	GAT	GAG	416
	Leu	Gly	Tyr	Phe	Tyr	Asn	Lys	Glu	Tyr	Gly	Lys	Ser	Val	Leu	Asp	Glu	
			70					75					80				
15	ACT	AAA	ACC	AAC	AAA	GAA	CTT	TCG	CCC	CTA	GTC	CCT	GCC	ACC	GGC	ACG	464
	Thr	Lys	Thr	Asn	Lys	Glu	Leu	Ser	Pro	Leu	Val	Pro	Ala	Thr	Gly	Thr	
		85					90					95					
	CTT	AAC	CCT	GCA	CTT	AAT	CCC	ACA	TTA	AAC	CCA	ACG	CTC	AAC	CCT	TTA	512
	Leu	Asn	Pro	Ala	Leu	Asn	Pro	Thr	Leu	Asn	Pro	Thr	Leu	Asn	Pro	Leu	
20	100					105					110				115		
	GAG	CAA	GCC	CCA	ACT	AAT	CCT	TTA	ATG	CCA	CAA	CAA	ACG	CCT	AAC	GAA	560
	Glu	Gln	Ala	Pro	Thr	Asn	Pro	Leu	Met	Pro	Gln	Gln	Thr	Pro	Asn	Glu	
					120					125					130		
	CTC	CCT	AAA	GAG	CCA	GCC	AAA	ACG	CCT	TCT	GTT	GAA	AGC	CCC	AAA	CAG	608
25	Leu	Pro	Lys	Glu	Pro	Ala	Lys	Thr	Pro	Ser	Val	Glu	Ser	Pro	Lys	Gln	
				135				140						145			
	AAT	GAA	AAG	AAT	GAA	AAG	AAT	GAC	GCC	AAA	GAG	AAT	GGT	ATA	AAG	GGT	656
	Asn	Glu	Lys	Asn	Glu	Lys	Asn	Asp	Ala	Lys	Glu	Asn	Gly	Ile	Lys	Gly	
			150					155					160				
30	GTT	GAA	AAA	ACC	AAA	GAG	AAC	GCC	AAA	ACG	CCC	CCA	ACC	ACC	CAC	CAA	704
	Val	Glu	Lys	Thr	Lys	Glu	Asn	Ala	Lys	Thr	Pro	Pro	Thr	Thr	His	Gln	
		165					170					175					
	AAG	CCT	AAA	ACG	CAT	GCA	ACG	CAA	ACC	AAC	GCC	CAT	ACC	AAC	CAA	AAA	752
	Lys	Pro	Lys	Thr	His	Ala	Thr	Gln	Thr	Asn	Ala	His	Thr	Asn	Gln	Lys	
35	180					185					190				195		
	AAG	GAT	GAA	AAA	TAATGTTACA	GGCCATTTAT	AACGAAACCA	AAGATCTGAT	GCAAA								809
	Lys	Asp	Glu	Lys													

AAAGCATTCA AGCTTTAAAC AGGGATTTTT CCACTCTAAG GAGCGCGAAA GTTTCAGTCA	869
ATATTTTAGA TCACATCAAA GTGGATTATT ACGGCACGCC CACGGCATTAA AATCAAGTCG	929
GATCCGTGAT GAGCTTGGAT GCGACCACCC TT	961

(2) INFORMATION FOR SEQ ID NO:40:

5 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 199 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

10 (ii) MOLECULE TYPE: protein
 (v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

	Met	Thr	Ser	Ala	Leu	Leu	Gly	Leu	Gln	Ile	Val	Leu	Ala	Val	Leu	Ile	
	1				5					10					15		
15	Val	Val	Val	Val	Leu	Leu	Gln	Lys	Ser	Ser	Ser	Ile	Gly	Leu	Gly	Ala	
				20				25						30			
	Tyr	Ser	Gly	Ser	Asn	Glu	Ser	Leu	Phe	Gly	Ala	Lys	Gly	Pro	Ala	Ser	
		35					40						45				
	Phe	Met	Ala	Lys	Leu	Thr	Met	Phe	Leu	Gly	Leu	Leu	Phe	Val	Ile	Asn	
20		50					55					60					
	Thr	Ile	Ala	Leu	Gly	Tyr	Phe	Tyr	Asn	Lys	Glu	Tyr	Gly	Lys	Ser	Val	
	65					70					75				80		
	Leu	Asp	Glu	Thr	Lys	Thr	Asn	Lys	Glu	Leu	Ser	Pro	Leu	Val	Pro	Ala	
				85						90				95			
25	Thr	Gly	Thr	Leu	Asn	Pro	Ala	Leu	Asn	Pro	Thr	Leu	Asn	Pro	Thr	Leu	
				100					105					110			
	Asn	Pro	Leu	Glu	Gln	Ala	Pro	Thr	Asn	Pro	Leu	Met	Pro	Gln	Gln	Thr	
		115					120						125				
	Pro	Asn	Glu	Leu	Pro	Lys	Glu	Pro	Ala	Lys	Thr	Pro	Ser	Val	Glu	Ser	
30		130					135						140				
	Pro	Lys	Gln	Asn	Glu	Lys	Asn	Glu	Lys	Asn	Asp	Ala	Lys	Glu	Asn	Gly	
	145				150					155					160		
	Ile	Lys	Gly	Val	Glu	Lys	Thr	Lys	Glu	Asn	Ala	Lys	Thr	Pro	Pro	Thr	
				165					170					175			
35	Thr	His	Gln	Lys	Pro	Lys	Thr	His	Ala	Thr	Gln	Thr	Asn	Ala	His	Thr	
				180				185						190			
	Asn	Gln	Lys	Lys	Asp	Glu	Lys										
				195													

(2) INFORMATION FOR SEQ ID NO:41:

40 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1058 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence

(B) LOCATION: 325...879

(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

```

10 CCTAGTCCCT GCCACCGGCA CGCTTAACCC TGCACCTAAT CCCACATTAA ACCCAACGCT 60
    CAACCCTTTA GAGCAAGCCC CAACTAATCC TTTAATGCCA CAACAAACGC CTAACGAACT 120
    CCCTAAAGAG CCAGCCAAAA CGCCTTCTGT TGAAAGCCCC AAACAGAATG AAAAGAATGA 180
    AAAGAATGAC GCCAAAGAGA ATGGTATAAA GGGTGTGAA AAAACCAAAG AGAACGCCAA 240
    AACGCCCCCA ACCACCCACC AAAAGCCTAA AACGCATGCA ACGCAAACCA ACGCCCATAC 300
    CAACCAAAAA AAGGATGAAA AATA ATG TTA CAG GCC ATT TAT AAC GAA ACC 351
                                Met Leu Gln Ala Ile Tyr Asn Glu Thr
15                                1                    5

    AAA GAT CTG ATG CAA AAA AGC ATT CAA GCT TTA AAC AGG GAT TTT TCC 399
    Lys Asp Leu Met Gln Lys Ser Ile Gln Ala Leu Asn Arg Asp Phe Ser
    10                    15                    20                    25

    ACT CTA AGG AGC GCG AAA GTT TCA GTC AAT ATT TTA GAT CAC ATC AAA 447
    Thr Leu Arg Ser Ala Lys Val Ser Val Asn Ile Leu Asp His Ile Lys
    20                    30                    35                    40

    GTG GAT TAT TAC GGC ACG CCC ACG GCA TTA AAT CAA GTC GGA TCC GTG 495
    Val Asp Tyr Tyr Gly Thr Pro Thr Ala Leu Asn Gln Val Gly Ser Val
    45                    50                    55

    25 ATG AGC TTG GAT GCG ACC ACC CTT CAA ATC AGC CCA TGG GAA AAA AAC 543
    Met Ser Leu Asp Ala Thr Thr Leu Gln Ile Ser Pro Trp Glu Lys Asn
    60                    65                    70

    CTG CTC AAA GAA ATT GAA AGA TCC ATT CAA GAA GCC AAT ATT GGT GTC 591
    Leu Leu Lys Glu Ile Glu Arg Ser Ile Gln Glu Ala Asn Ile Gly Val
    30                    75                    80                    85

    AAT CCT AAT AAC GAC GGC GAA ACG ATC AAG CTT TTT TTC CCG CCC ATG 639
    Asn Pro Asn Asn Asp Gly Glu Thr Ile Lys Leu Phe Phe Pro Pro Met
    90                    95                    100                    105

    ACA AGT GAG CAA AGA AAA CTC ATC GCA AAA GAC GCC AAA GCG ATG GGT 687
    Thr Ser Glu Gln Arg Lys Leu Ile Ala Lys Asp Ala Lys Ala Met Gly
    35                    110                    115                    120

    GAA AAG GCT AAA GTG GCT GTG AGG AAT ATC CGC CAA GAT GCT AAC AAC 735
    Glu Lys Ala Lys Val Ala Val Arg Asn Ile Arg Gln Asp Ala Asn Asn
```

	125	130	135	
	CAG GTG AAA AAA TTA GAA AAA GAC AAA GAA ATC AGC GAA GAT GAA AGC			783
	Gln Val Lys Lys Leu Glu Lys Asp Lys Glu Ile Ser Glu Asp Glu Ser			
	140	145	150	
5	AAA AAA GCC CAA GAG CAG ATC CAA AAA ATC ACC GAT GAA GCC ATT AAA			831
	Lys Lys Ala Gln Glu Gln Ile Gln Lys Ile Thr Asp Glu Ala Ile Lys			
	155	160	165	
	AAA ATT GAT GAA AGC GTG AAA AAC AAA GAA GAC GCG ATC TTA AAG GTC T			880
	Lys Ile Asp Glu Ser Val Lys Asn Lys Glu Asp Ala Ile Leu Lys Val			
10	170	175	180	185
	AAACCATGGA TATTAAGGCA TGTTATCAAA ACGCTAAAGC GTTATTAGAG GGGCATTCT			940
	TGCTCAGCAG TGGGTTTCAT TCCAATTATT ATTTGCAATC CGCTAAAGTT TTAGAAGATC			1000
	CCAAACTAGC CGAACAATTA GCGCTAGAAT TAGCCAAACA AATCCAAGAA GCTCATTT			1058

(2) INFORMATION FOR SEQ ID NO:42:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 185 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

20 (ii) MOLECULE TYPE: protein
 (v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

	Met	Leu	Gln	Ala	Ile	Tyr	Asn	Glu	Thr	Lys	Asp	Leu	Met	Gln	Lys	Ser
	1				5					10					15	
25	Ile	Gln	Ala	Leu	Asn	Arg	Asp	Phe	Ser	Thr	Leu	Arg	Ser	Ala	Lys	Val
				20						25					30	
	Ser	Val	Asn	Ile	Leu	Asp	His	Ile	Lys	Val	Asp	Tyr	Tyr	Gly	Thr	Pro
				35						40					45	
	Thr	Ala	Leu	Asn	Gln	Val	Gly	Ser	Val	Met	Ser	Leu	Asp	Ala	Thr	Thr
30				50						55					60	
	Leu	Gln	Ile	Ser	Pro	Trp	Glu	Lys	Asn	Leu	Leu	Lys	Glu	Ile	Glu	Arg
	65					70						75				80
	Ser	Ile	Gln	Glu	Ala	Asn	Ile	Gly	Val	Asn	Pro	Asn	Asn	Asp	Gly	Glu
					85						90				95	
35	Thr	Ile	Lys	Leu	Phe	Phe	Pro	Pro	Met	Thr	Ser	Glu	Gln	Arg	Lys	Leu
				100						105					110	
	Ile	Ala	Lys	Asp	Ala	Lys	Ala	Met	Gly	Glu	Lys	Ala	Lys	Val	Ala	Val
				115					120					125		
	Arg	Asn	Ile	Arg	Gln	Asp	Ala	Asn	Asn	Gln	Val	Lys	Lys	Leu	Glu	Lys
40				130										140		
	Asp	Lys	Glu	Ile	Ser	Glu	Asp	Glu	Ser	Lys	Lys	Ala	Gln	Glu	Gln	Ile

145	150	155	160
Gln Lys Ile Thr Asp Glu Ala Ile Lys Lys Ile Asp Glu Ser Val Lys			
	165	170	175
Asn Lys Glu Asp Ala Ile Leu Lys Val			
5	180	185	

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 1669 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ix) FEATURE:

- 15 (A) NAME/KEY: Coding Sequence
 (B) LOCATION: 163...1389
 (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

20	GAGTGGATGA AAAAGACACT TTCAATTTTG CAAAATTGG CTATGAACAG GGCAAGGGCG	60
	AAGAATTAAA AGAAGTAGAA GAAAAGCATG CGTTTAAGAA AATCCCTTTT GTCAAAGATT	120
	TGCACAAAAT CGCCCCCACT ATCTTAAAAA AGAGGCTATA AA ATG GCT CAA AAT	174
	Met Ala Gln Asn	
	1	
25	TTC ACG AAA CTC AAC CCC CAG TTT GAA AAC ATC ATT TTT GAA CAT GAC	222
	Phe Thr Lys Leu Asn Pro Gln Phe Glu Asn Ile Ile Phe Glu His Asp	
	5 10 15 20	
	GAC AAC CAA ATG ATT TTA AAC TTT GGC CCC CAA CAC CCC AGT AGT CAT	270
	Asp Asn Gln Met Ile Leu Asn Phe Gly Pro Gln His Pro Ser Ser His	
	25 30 35	
30	GGG CAA TTG CGC TTG ATT TTG GAA TTA GAG GGC GAA AAA ATC ATT AAG	318
	Gly Gln Leu Arg Leu Ile Leu Glu Leu Glu Gly Glu Lys Ile Ile Lys	
	40 45 50	
	GCT ACC CCT GAA ATT GGC TAC TTG CAT AGA GGC TGT GAA AAG TTA GGC	366
	Ala Thr Pro Glu Ile Gly Tyr Leu His Arg Gly Cys Glu Lys Leu Gly	
	55 60 65	
35	GAA AAC ATG ACC TAT AAC GAA TAC ATG CCC ACT ACT GAT AGA TTG GAT	414
	Glu Asn Met Thr Tyr Asn Glu Tyr Met Pro Thr Thr Asp Arg Leu Asp	
	70 75 80	
	TAC ACT TCT TCT ACC AGC AAT AAT TAC GCT TAC GCT TAT GCG GTA GAG	462

	Tyr	Thr	Ser	Ser	Thr	Ser	Asn	Asn	Tyr	Ala	Tyr	Ala	Tyr	Ala	Val	Glu	
	85					90					95					100	
	ACC	TTA	CTC	AAT	TTA	GAA	ATC	CCA	CGC	CGA	GCG	CAG	GTG	ATC	CGC	ACG	510
5	Thr	Leu	Leu	Asn	Leu	Glu	Ile	Pro	Arg	Arg	Ala	Gln	Val	Ile	Arg	Thr	
					105					110					115		
	ATT	TTA	CTA	GAG	CTT	AAC	CGC	ATG	ATC	TCA	CAC	ATC	TTT	TTT	ATC	AGC	558
	Ile	Leu	Leu	Glu	Leu	Asn	Arg	Met	Ile	Ser	His	Ile	Phe	Phe	Ile	Ser	
				120					125					130			
	GTG	CAT	GCT	TTA	GAT	GTG	GGG	GCG	ATG	AGC	GTG	TTT	TTG	TAT	GCG	TTT	606
10	Val	His	Ala	Leu	Asp	Val	Gly	Ala	Met	Ser	Val	Phe	Leu	Tyr	Ala	Phe	
			135					140					145				
	AAA	ACG	AGG	GAA	TAC	GGC	TTG	GAT	TTG	ATG	GAG	GAT	TAT	TGC	GGG	GCT	654
	Lys	Thr	Arg	Glu	Tyr	Gly	Leu	Asp	Leu	Met	Glu	Asp	Tyr	Cys	Gly	Ala	
		150					155					160					
15	AGG	CTC	ACG	CAT	AAC	GCT	ATA	AGG	ATT	GGG	GGC	GTG	CCT	TTA	GAT	TTA	702
	Arg	Leu	Thr	His	Asn	Ala	Ile	Arg	Ile	Gly	Gly	Val	Pro	Leu	Asp	Leu	
	165					170					175					180	
	CCC	CCT	AAT	TGG	TTA	GAA	GGC	TTA	AAA	AAG	TTT	TTA	GGC	GAA	ATG	AGG	750
20	Pro	Pro	Asn	Trp	Leu	Glu	Gly	Leu	Lys	Lys	Phe	Leu	Gly	Glu	Met	Arg	
					185					190					195		
	GAA	TGC	AAA	AAA	CTC	ATT	CAA	GGC	TTA	TTG	GAT	AAG	AAT	CGC	ATT	TGG	798
	Glu	Cys	Lys	Lys	Leu	Ile	Gln	Gly	Leu	Leu	Asp	Lys	Asn	Arg	Ile	Trp	
				200					205					210			
	CGG	ATG	CGC	TTG	GAA	AAT	GTG	GGC	GTT	GTA	ACG	CAA	AAA	ATG	GCG	CAA	846
25	Arg	Met	Arg	Leu	Glu	Asn	Val	Gly	Val	Val	Thr	Gln	Lys	Met	Ala	Gln	
			215					220					225				
	AGC	TGG	GGC	ATG	AGC	GGT	ATC	ATG	TTA	AGA	GGG	ACT	GGG	ATC	GCT	TAT	894
	Ser	Trp	Gly	Met	Ser	Gly	Ile	Met	Leu	Arg	Gly	Thr	Gly	Ile	Ala	Tyr	
		230					235					240					
30	GAC	ATC	AGA	AAA	GAA	GAG	CCT	TAT	GAG	CTT	TAT	AAA	GAG	CTT	GAT	TTT	942
	Asp	Ile	Arg	Lys	Glu	Glu	Pro	Tyr	Glu	Leu	Tyr	Lys	Glu	Leu	Asp	Phe	
	245					250					255					260	
	GAT	GTG	CCG	GTG	GGC	AAT	TAT	GGC	GAT	AGT	TAT	GAT	AGG	TAT	TGT	TTG	990
35	Asp	Val	Pro	Val	Gly	Asn	Tyr	Gly	Asp	Ser	Tyr	Asp	Arg	Tyr	Cys	Leu	
					265					270					275		
	TAT	ATG	TTA	GAA	ATT	GAT	GAA	AGC	GTT	CGC	ATC	ATT	GAA	CAG	CTC	ATT	1038
	Tyr	Met	Leu	Glu	Ile	Asp	Glu	Ser	Val	Arg	Ile	Ile	Glu	Gln	Leu	Ile	
				280					285					290			

	CCT ATG TAT GCT AAA ACC GAT ACG CCT ATC ATG GCT CAA AAC CCG CAT	1086
	Pro Met Tyr Ala Lys Thr Asp Thr Pro Ile Met Ala Gln Asn Pro His	
	295 300 305	
	TAT ATT TCC GCC CCT AAA GAA GAT ATA ATG ACG CAA AAC TAC GCC TTG	1134
5	Tyr Ile Ser Ala Pro Lys Glu Asp Ile Met Thr Gln Asn Tyr Ala Leu	
	310 315 320	
	ATG CAG CAT TTT GTT TTA GTG GCT CAG GGC ATG CGT CCG CCC GTT GGG	1182
	Met Gln His Phe Val Leu Val Ala Gln Gly Met Arg Pro Pro Val Gly	
	325 330 335 340	
10	GAA GTG TAT GCC CCC ACA GAA AGC CCT AAA GGG GAA TTA GGG TTT TTT	1230
	Glu Val Tyr Ala Pro Thr Glu Ser Pro Lys Gly Glu Leu Gly Phe Phe	
	345 350 355	
	ATC CAT TCA GAG GGC GAG CCT TAC CCT CAC AGG CTA AAA ATC AGA GCC	1278
15	Ile His Ser Glu Gly Glu Pro Tyr Pro His Arg Leu Lys Ile Arg Ala	
	360 365 370	
	CCT AGC TTT TAT CAC ATT GGG GCT TTG AGC GAC ATT TTA GTG GGG CAA	1326
	Pro Ser Phe Tyr His Ile Gly Ala Leu Ser Asp Ile Leu Val Gly Gln	
	375 380 385	
	TAT TTA GCG GAT GCA GTA ACC GTG ATT GGC TCA ACC AAT GCG GTG TTT	1374
20	Tyr Leu Ala Asp Ala Val Thr Val Ile Gly Ser Thr Asn Ala Val Phe	
	390 395 400	
	GGC GAG GTG GAT AGA TGAAACGCTT TGATTACGC CCCTTAAAAG CGGGTATTTT T	1430
	Gly Glu Val Asp Arg	
	405	
25	GAACGCTTAG AAGAATTGAT TGAAAAAGAA ATGCAACCTA ATGAAGTCGC TATTTTCATG	1490
	TTTGAAGTGG GGGATTTTTC TAATATCCCT AAGAGCGCTG AATTTATCCA ATCTAAAGGG	1550
	CATGAGCTCC TCAATTCTTT GCGTTTCAAT CAAGCGGATT GGACGATTGT CGTGAGAAAA	1610
	AAGGCTTGAT TTTGAGCGGC TTTAACCCTT TAAATCTCC CTTAGTCGCA AGCTCTTCT	1669

(2) INFORMATION FOR SEQ ID NO:44:

30 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 409 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

35 (ii) MOLECULE TYPE: protein
 (v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

	Met	Ala	Gln	Asn	Phe	Thr	Lys	Leu	Asn	Pro	Gln	Phe	Glu	Asn	Ile	Ile	
	1				5				10						15		
	Phe	Glu	His	Asp	Asp	Asn	Gln	Met	Ile	Leu	Asn	Phe	Gly	Pro	Gln	His	
				20					25					30			
5	Pro	Ser	Ser	His	Gly	Gln	Leu	Arg	Leu	Ile	Leu	Glu	Leu	Glu	Gly	Glu	
			35				40					45					
	Lys	Ile	Ile	Lys	Ala	Thr	Pro	Glu	Ile	Gly	Tyr	Leu	His	Arg	Gly	Cys	
		50				55					60						
	Glu	Lys	Leu	Gly	Glu	Asn	Met	Thr	Tyr	Asn	Glu	Tyr	Met	Pro	Thr	Thr	
10	65					70				75						80	
	Asp	Arg	Leu	Asp	Tyr	Thr	Ser	Ser	Thr	Ser	Asn	Asn	Tyr	Ala	Tyr	Ala	
				85					90					95			
	Tyr	Ala	Val	Glu	Thr	Leu	Leu	Asn	Leu	Glu	Ile	Pro	Arg	Arg	Ala	Gln	
				100					105					110			
15	Val	Ile	Arg	Thr	Ile	Leu	Leu	Glu	Leu	Asn	Arg	Met	Ile	Ser	His	Ile	
			115					120					125				
	Phe	Phe	Ile	Ser	Val	His	Ala	Leu	Asp	Val	Gly	Ala	Met	Ser	Val	Phe	
		130					135					140					
	Leu	Tyr	Ala	Phe	Lys	Thr	Arg	Glu	Tyr	Gly	Leu	Asp	Leu	Met	Glu	Asp	
20	145					150					155					160	
	Tyr	Cys	Gly	Ala	Arg	Leu	Thr	His	Asn	Ala	Ile	Arg	Ile	Gly	Gly	Val	
				165					170						175		
	Pro	Leu	Asp	Leu	Pro	Pro	Asn	Trp	Leu	Glu	Gly	Leu	Lys	Lys	Phe	Leu	
				180					185					190			
25	Gly	Glu	Met	Arg	Glu	Cys	Lys	Lys	Leu	Ile	Gln	Gly	Leu	Leu	Asp	Lys	
			195					200					205				
	Asn	Arg	Ile	Trp	Arg	Met	Arg	Leu	Glu	Asn	Val	Gly	Val	Val	Thr	Gln	
		210					215					220					
	Lys	Met	Ala	Gln	Ser	Trp	Gly	Met	Ser	Gly	Ile	Met	Leu	Arg	Gly	Thr	
30	225					230					235					240	
	Gly	Ile	Ala	Tyr	Asp	Ile	Arg	Lys	Glu	Glu	Pro	Tyr	Glu	Leu	Tyr	Lys	
				245						250					255		
	Glu	Leu	Asp	Phe	Asp	Val	Pro	Val	Gly	Asn	Tyr	Gly	Asp	Ser	Tyr	Asp	
			260						265				270				
35	Arg	Tyr	Cys	Leu	Tyr	Met	Leu	Glu	Ile	Asp	Glu	Ser	Val	Arg	Ile	Ile	
			275					280					285				
	Glu	Gln	Leu	Ile	Pro	Met	Tyr	Ala	Lys	Thr	Asp	Thr	Pro	Ile	Met	Ala	
		290					295					300					
	Gln	Asn	Pro	His	Tyr	Ile	Ser	Ala	Pro	Lys	Glu	Asp	Ile	Met	Thr	Gln	
40	305					310					315					320	
	Asn	Tyr	Ala	Leu	Met	Gln	His	Phe	Val	Leu	Val	Ala	Gln	Gly	Met	Arg	
				325						330					335		
	Pro	Pro	Val	Gly	Glu	Val	Tyr	Ala	Pro	Thr	Glu	Ser	Pro	Lys	Gly	Glu	
			340						345				350				
45	Leu	Gly	Phe	Phe	Ile	His	Ser	Glu	Gly	Glu	Pro	Tyr	Pro	His	Arg	Leu	
			355					360					365				
	Lys	Ile	Arg	Ala	Pro	Ser	Phe	Tyr	His	Ile	Gly	Ala	Leu	Ser	Asp	Ile	
		370					375					380					
	Leu	Val	Gly	Gln	Tyr	Leu	Ala	Asp	Ala	Val	Thr	Val	Ile	Gly	Ser	Thr	
50	385					390					395					400	

Asn Ala Val Phe Gly Glu Val Asp Arg
405

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 869 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:

- 10 (A) NAME/KEY: Coding Sequence
(B) LOCATION: 358...732
(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

15 TAACTTGTGG TTAACTACCG CCAGACTCCT TTTGAGTTTG GCAAACGCGC CAATGAGTTC 60
TTTAGGCATT TTTTCAGTGC CGATCTTAAA GTTTTCAAGA CTGCGTTGCG TTTGAGCCCC 120
CCAATATTGG CTATCATTTA CTTTGATTTT GCCCATCGTG TCATGTTCAA TTCTAAATTG 180
CATGCTAATC CTTTGAAATT TGATTTTAAA ACCTTAAAAA AATAGCATAA ACTCTTATAC 240
CTTCTACTTA AAAACCCTAA TTTTAAAC ACCATTCCA CAATTTTAC ACAAAGAGG 300
20 GTTATTATCC GTTCGCAACA AGAATTTTCT TGTATCTTA ATGTAAAGGT CAAAACG ATG 360
Met
1
AAA AAG TTA GCC GCT TTA TTT TTA GTA AGC GTG TTG GGG GTT ATG GGT 408
Lys Lys Leu Ala Ala Leu Phe Leu Val Ser Val Leu Gly Val Met Gly
25 5 10 15
TTA AAC GCA TGG GAG CAA ACC CTA AAA GCT AAT GAC TTG GAA GTG AAA 456
Leu Asn Ala Trp Glu Gln Thr Leu Lys Ala Asn Asp Leu Glu Val Lys
20 25 30
ATC AAA TCC GTG GGT AAC CCC ATT AAA GGC GAT AAC ACT TTC ATT CTC 504
30 Ile Lys Ser Val Gly Asn Pro Ile Lys Gly Asp Asn Thr Phe Ile Leu
35 40 45
AGC CCC ACT TTA AAA GGT AAG GCT TTA GAA AAA GCT ATC GTT AGG GTG 552
Ser Pro Thr Leu Lys Gly Lys Ala Leu Glu Lys Ala Ile Val Arg Val
50 55 60 65
35 CAG TTT ATG ATG CCT GAA ATG CCC GGC ATG CCA GCG ATG AAA GAA ATG 600
Gln Phe Met Met Pro Glu Met Pro Gly Met Pro Ala Met Lys Glu Met
70 75 80
GCG CAA GTG AGT GAA AAA AAC GGC CTT TAT GAA GCT AAA ACC AAT CTT 648

Ala Gln Val Ser Glu Lys Asn Gly Leu Tyr Glu Ala Lys Thr Asn Leu
85 90 95

TCT ATG AAC GGG ACA TGG CAG GTT AGG GTG GAT ATT AAA TCT AAA GAG 696
Ser Met Asn Gly Thr Trp Gln Val Arg Val Asp Ile Lys Ser Lys Glu
5 100 105 110

GGT CAG GTT TAT CGC GCT AAA ACA AGC CTG GAT TTA TAAGAGCATG CTATCT 748
Gly Gln Val Tyr Arg Ala Lys Thr Ser Leu Asp Leu
115 120 125

TTTATAAGCG CGTTTGATAA AAGGGGCGTT TCAATACGCC TTTTAACAGC CTTGTTACTG 808
10 CTTTTTAGTT TGGGTTTGGC TAAAGATTTA GAGATCCAAT CTTTTGTGGC TAAATACCTT 868
T 869

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 125 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

Met Lys Lys Leu Ala Ala Leu Phe Leu Val Ser Val Leu Gly Val Met
1 5 10 15
Gly Leu Asn Ala Trp Glu Gln Thr Leu Lys Ala Asn Asp Leu Glu Val
20 25 30
Lys Ile Lys Ser Val Gly Asn Pro Ile Lys Gly Asp Asn Thr Phe Ile
35 40 45
Leu Ser Pro Thr Leu Lys Gly Lys Ala Leu Glu Lys Ala Ile Val Arg
50 55 60
Val Gln Phe Met Met Pro Glu Met Pro Gly Met Pro Ala Met Lys Glu
65 70 75 80
Met Ala Gln Val Ser Glu Lys Asn Gly Leu Tyr Glu Ala Lys Thr Asn
85 90 95
Leu Ser Met Asn Gly Thr Trp Gln Val Arg Val Asp Ile Lys Ser Lys
100 105 110
35 Glu Gly Gln Val Tyr Arg Ala Lys Thr Ser Leu Asp Leu
115 120 125

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1217 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:

5 (A) NAME/KEY: Coding Sequence
(B) LOCATION: 73...1152
(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

10	TCCATGCGTT TTGATGCGAT TTTAAAAAAT CTTTGGGTAT TTTAGCATGC CAATGGTTAA	60
	AAAAAGGTGG TT ATG AAT GGT TTT TGC GCT AGA CTA CGA GCC ATA ACT CAT	111
	Met Asn Gly Phe Cys Ala Arg Leu Arg Ala Ile Thr His	
	1 5 10	
15	AAT GAA AGA TTA AAA ATG AAA ATA GCG GTA TTA CTC AGT GGG GGG GTG	159
	Asn Glu Arg Leu Lys Met Lys Ile Ala Val Leu Leu Ser Gly Gly Val	
	15 20 25	
	GAT AGC TCT TAT AGC GCT TAT AGC TTA AAA GAG CAA GGG CAT GAA TTA	207
	Asp Ser Ser Tyr Ser Ala Tyr Ser Leu Lys Glu Gln Gly His Glu Leu	
	30 35 40 45	
20	GTG GGG ATT TAT TTA AAA CTC CAT GCG AGT GAA AAA AAG CAT GAT TTA	255
	Val Gly Ile Tyr Leu Lys Leu His Ala Ser Glu Lys Lys His Asp Leu	
	50 55 60	
	TAC ATC AAA AAC GCT CAA AAA GCA TGC GAG TTT TTA GGC ATT CCT TTA	303
	Tyr Ile Lys Asn Ala Gln Lys Ala Cys Glu Phe Leu Gly Ile Pro Leu	
	65 70 75	
25	GAG GTG TTG GAT TTT CAA AAG GAT TTT AAA AGC GCG GTT TAT GAT GAA	351
	Glu Val Leu Asp Phe Gln Lys Asp Phe Lys Ser Ala Val Tyr Asp Glu	
	80 85 90	
	TTT ATC AAC GCC TAT GAA GAA GGG CAA ACC CCA AAC CCT TGT GCG TTG	399
30	Phe Ile Asn Ala Tyr Glu Glu Gly Gln Thr Pro Asn Pro Cys Ala Leu	
	95 100 105	
	TGC AAC CCT TTA ATG AAG TTT GGG CTA GCT TTG GAT CAC GCT TTA AAA	447
	Cys Asn Pro Leu Met Lys Phe Gly Leu Ala Leu Asp His Ala Leu Lys	
	110 115 120 125	
35	TTA GGG TGT GAA AAG ATC GCT ACC GGG CAT TAT GCG AGA GTC AAA GAA	495
	Leu Gly Cys Glu Lys Ile Ala Thr Gly His Tyr Ala Arg Val Lys Glu	
	130 135 140	
	ATT GAC AAA ATA AGT TAT ATT CAA GAG GCT TTG GAT AAA ACT AAA GAT	543

	Ile	Asp	Lys	Ile	Ser	Tyr	Ile	Gln	Glu	Ala	Leu	Asp	Lys	Thr	Lys	Asp	
				145					150					155			
	CAG	AGC	TAT	TTT	TTA	TAC	GCT	TTA	GAG	CAT	GAA	GTG	ATC	GCT	AAA	TTG	591
5	Gln	Ser	Tyr	Phe	Leu	Tyr	Ala	Leu	Glu	His	Glu	Val	Ile	Ala	Lys	Leu	
			160					165					170				
	GTG	TTC	CCT	TTA	GGG	GAT	TTG	CTA	AAA	AAG	GAT	ATT	AAG	CCT	TTA	GCC	639
	Val	Phe	Pro	Leu	Gly	Asp	Leu	Leu	Lys	Lys	Asp	Ile	Lys	Pro	Leu	Ala	
			175				180					185					
10	TTG	AAT	GCG	ATG	CCT	TTT	TTA	GGC	ACT	TTA	GAG	ACT	TAT	AAG	GAA	TCT	687
	Leu	Asn	Ala	Met	Pro	Phe	Leu	Gly	Thr	Leu	Glu	Thr	Tyr	Lys	Glu	Ser	
						195					200					205	
	CAA	GAA	ATC	TGC	TTT	GTG	GAA	AAA	AGC	TAC	ATT	GAC	ACT	TTA	AAA	AAG	735
	Gln	Glu	Ile	Cys	Phe	Val	Glu	Lys	Ser	Tyr	Ile	Asp	Thr	Leu	Lys	Lys	
				210						215					220		
15	CAT	GTT	GAA	GTG	GAA	AAA	GAG	GGC	GTG	GTG	AAA	AAC	CTA	CAA	GGC	GAA	783
	His	Val	Glu	Val	Glu	Lys	Glu	Gly	Val	Val	Lys	Asn	Leu	Gln	Gly	Glu	
				225				230						235			
	GTC	ATT	GGC	ACG	CAT	AAA	GGC	TAT	ATG	CAA	TAC	ACG	ATT	GGC	AAA	CGC	831
20	Val	Ile	Gly	Thr	His	Lys	Gly	Tyr	Met	Gln	Tyr	Thr	Ile	Gly	Lys	Arg	
			240					245					250				
	AAA	GGC	TTT	AGT	ATT	AAA	GGC	GCG	TTA	GAG	CCG	CAT	TTT	GTG	GTG	GGG	879
	Lys	Gly	Phe	Ser	Ile	Lys	Gly	Ala	Leu	Glu	Pro	His	Phe	Val	Val	Gly	
			255				260					265					
25	ATT	GAC	GCT	AAA	AAG	AAC	GAG	CTA	GTC	GTG	GGC	AAA	AAA	GAA	GAT	CTC	927
	Ile	Asp	Ala	Lys	Lys	Asn	Glu	Leu	Val	Val	Gly	Lys	Lys	Glu	Asp	Leu	
						275					280					285	
	GCC	ACG	CAT	TCG	CTT	AAG	GCT	AAA	AAC	AAA	TCT	TTA	ATG	AAA	GAT	TTT	975
	Ala	Thr	His	Ser	Leu	Lys	Ala	Lys	Asn	Lys	Ser	Leu	Met	Lys	Asp	Phe	
					290					295					300		
30	AAA	GAT	GGC	GAA	TAT	TTT	ATC	AAG	GCT	CGT	TAC	AGG	AGC	GTG	CCT	GCT	1023
	Lys	Asp	Gly	Glu	Tyr	Phe	Ile	Lys	Ala	Arg	Tyr	Arg	Ser	Val	Pro	Ala	
				305					310					315			
	AAA	GCG	CAT	GTG	AGT	TTG	AAA	GAT	GAG	GTG	ATT	GAA	GTG	GGG	TTT	AAA	1071
35	Lys	Ala	His	Val	Ser	Leu	Lys	Asp	Glu	Val	Ile	Glu	Val	Gly	Phe	Lys	
			320					325					330				
	GAG	CCT	TTT	TAT	GGC	GTG	GCT	AAA	GGG	CAA	GCT	TTG	GTC	GTT	TAT	AAA	1119
	Glu	Pro	Phe	Tyr	Gly	Val	Ala	Lys	Gly	Gln	Ala	Leu	Val	Val	Tyr	Lys	
			335				340					345					

GAT GAC ATC TTG CTT GGT GGG GGC GTG ATT GTT TAAAACTAA AGAACTAAGA 1172
 Asp Asp Ile Leu Leu Gly Gly Val Ile Val
 350 355 360

GATACGCCTT TTGGCAGTCT CTTAATGTTT TATTGAATAG GCGTT 1217

5 (2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 360 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

15	Met	Asn	Gly	Phe	Cys	Ala	Arg	Leu	Arg	Ala	Ile	Thr	His	Asn	Glu	Arg
	1				5					10					15	
	Leu	Lys	Met	Lys	Ile	Ala	Val	Leu	Leu	Ser	Gly	Gly	Val	Asp	Ser	Ser
				20					25					30		
	Tyr	Ser	Ala	Tyr	Ser	Leu	Lys	Glu	Gln	Gly	His	Glu	Leu	Val	Gly	Ile
			35					40					45			
20	Tyr	Leu	Lys	Leu	His	Ala	Ser	Glu	Lys	Lys	His	Asp	Leu	Tyr	Ile	Lys
		50					55					60				
	Asn	Ala	Gln	Lys	Ala	Cys	Glu	Phe	Leu	Gly	Ile	Pro	Leu	Glu	Val	Leu
	65					70					75					80
	Asp	Phe	Gln	Lys	Asp	Phe	Lys	Ser	Ala	Val	Tyr	Asp	Glu	Phe	Ile	Asn
25					85					90					95	
	Ala	Tyr	Glu	Glu	Gly	Gln	Thr	Pro	Asn	Pro	Cys	Ala	Leu	Cys	Asn	Pro
				100					105					110		
	Leu	Met	Lys	Phe	Gly	Leu	Ala	Leu	Asp	His	Ala	Leu	Lys	Leu	Gly	Cys
			115					120					125			
30	Glu	Lys	Ile	Ala	Thr	Gly	His	Tyr	Ala	Arg	Val	Lys	Glu	Ile	Asp	Lys
		130					135					140				
	Ile	Ser	Tyr	Ile	Gln	Glu	Ala	Leu	Asp	Lys	Thr	Lys	Asp	Gln	Ser	Tyr
	145					150					155				160	
	Phe	Leu	Tyr	Ala	Leu	Glu	His	Glu	Val	Ile	Ala	Lys	Leu	Val	Phe	Pro
35					165					170					175	
	Leu	Gly	Asp	Leu	Leu	Lys	Lys	Asp	Ile	Lys	Pro	Leu	Ala	Leu	Asn	Ala
				180				185					190			
	Met	Pro	Phe	Leu	Gly	Thr	Leu	Glu	Thr	Tyr	Lys	Glu	Ser	Gln	Glu	Ile
			195				200					205				
40	Cys	Phe	Val	Glu	Lys	Ser	Tyr	Ile	Asp	Thr	Leu	Lys	Lys	His	Val	Glu
		210					215					220				
	Val	Glu	Lys	Glu	Gly	Val	Val	Lys	Asn	Leu	Gln	Gly	Glu	Val	Ile	Gly
	225					230					235				240	
	Thr	His	Lys	Gly	Tyr	Met	Gln	Tyr	Thr	Ile	Gly	Lys	Arg	Lys	Gly	Phe

				245					250					255		
	Ser	Ile	Lys	Gly	Ala	Leu	Glu	Pro	His	Phe	Val	Val	Gly	Ile	Asp	Ala
				260					265					270		
	Lys	Lys	Asn	Glu	Leu	Val	Val	Gly	Lys	Lys	Glu	Asp	Leu	Ala	Thr	His
5			275					280					285			
	Ser	Leu	Lys	Ala	Lys	Asn	Lys	Ser	Leu	Met	Lys	Asp	Phe	Lys	Asp	Gly
		290				295					300					
	Glu	Tyr	Phe	Ile	Lys	Ala	Arg	Tyr	Arg	Ser	Val	Pro	Ala	Lys	Ala	His
	305				310					315				320		
10	Val	Ser	Leu	Lys	Asp	Glu	Val	Ile	Glu	Val	Gly	Phe	Lys	Glu	Pro	Phe
				325					330					335		
	Tyr	Gly	Val	Ala	Lys	Gly	Gln	Ala	Leu	Val	Val	Tyr	Lys	Asp	Asp	Ile
			340				345					350				
	Leu	Leu	Gly	Gly	Gly	Val	Ile	Val								
15			355				360									

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 975 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
 (B) LOCATION: 191...793
 (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

	ACATTACACA	TATCTGTCGC	TAAAACGCGC	CGCTTCACTA	AACCCACTGA	TTGTAAAAAT	60
	TTGTCTATTC	GCATGCGTTT	ATTTTACCCT	ATTCCTTAAG	TTTTTATCCA	TAACCTATAA	120
30	GGGTTTTAGT	TTTAGCATGT	TAGCATTGAG	CCACCACTCT	TTTAAAGGAA	TTTGTTTGAA	180
	GTTTCAAATT	ATG AGT TTG	TTA GCC ACT	CTT TTA TTA	GCC TCT TGC	TTG	229
		Met Ser Leu	Leu Ala Thr	Leu Leu Leu	Ala Ser Cys	Leu	
		1	5		10		
	CCC CCC AAA	GGC CAT CAT	TCT GGT TTG	GTG AAT CTT	TAT ATC GCT	CAT	277
35	Pro Pro Lys	Gly His His	Ser Gly Leu	Val Asn Leu	Tyr Ile Ala	His	
	15		20		25		
	CAA GGC CAA	AGC GTG CGC	ACT TAT TGG	CGC AAA GTG	GAT AGA GGA	GTT	325
	Gln Gly Gln	Ser Val Arg	Thr Tyr Trp	Arg Lys Val	Asp Arg Gly	Val	
	30		35		40	45	
40	ATC GCT AAA	CAC AAT GAA	GCG CTT AAA	AAA GAT CCT	AAA GCA AAG	CTC	373
	Ile Ala Lys	His Asn Glu	Ala Leu Lys	Lys Asp Pro	Lys Ala Lys	Leu	

	50	55	60	
	AAA GAC CCC AGG GGG CCT TTA TTC ATG CTA GGG AGT GAG CGC TTC ATG			421
	Lys Asp Pro Arg Gly Pro Leu Phe Met Leu Gly Ser Glu Arg Phe Met			
	65	70	75	
5	CTT TTA TGG AAA AAC CGC TAC GCT TTA GCC AAG CCC CAA TCG TTC AGG			469
	Leu Leu Trp Lys Asn Arg Tyr Ala Leu Ala Lys Pro Gln Ser Phe Arg			
	80	85	90	
	CTA GAG CCT GGT TTT TAT TAC TTG GAT TCT TTT AGC GTG GAA ACT CAA			517
	Leu Glu Pro Gly Phe Tyr Tyr Leu Asp Ser Phe Ser Val Glu Thr Gln			
10	95	100	105	
	AAA GGC GTC TTG CAG AGC GCT CCT GGC TAT TCA TAT ACT AAA AAT GGC			565
	Lys Gly Val Leu Gln Ser Ala Pro Gly Tyr Ser Tyr Thr Lys Asn Gly			
	110	115	120	125
	TAT GAT TTC AAA AAC AAC CGC CCC TTT TTC CTG GCC TTT GAA GTC AAA			613
15	Tyr Asp Phe Lys Asn Asn Arg Pro Phe Phe Leu Ala Phe Glu Val Lys			
	130	135	140	
	CCT GAT GGC AAA ACC ATT CTT CCT AGC GTG GAA TTA AGC CTG ATT AAA			661
	Pro Asp Gly Lys Thr Ile Leu Pro Ser Val Glu Leu Ser Leu Ile Lys			
	145	150	155	
20	ACC CCT AGA GGC TTT TTA GGG GTG TTC TTG TTT GAT AAT AAT GAA AAG			709
	Thr Pro Arg Gly Phe Leu Gly Val Phe Leu Phe Asp Asn Asn Glu Lys			
	160	165	170	
	GGG ACT AAC GCC AAG TGG ATT GAG GGG AGT TTG AAT TTA AAG CTT AAA			757
	Gly Thr Asn Ala Lys Trp Ile Glu Gly Ser Leu Asn Leu Lys Leu Lys			
25	175	180	185	
	AAC GCT TCC TTT AAA GAT GCG TGG GGG TTG GAA CAA TAAAGCATGA AGTGAT			809
	Asn Ala Ser Phe Lys Asp Ala Trp Gly Leu Glu Gln			
	190	195	200	
	CGCTTGCTTT TCGTAAGCTC TTTATGATTA GATTGTAAAA AAATGCCTTG AGTATTTTTT			869
30	AGATTTTATT ACCCCTATTC AATTGGAACA AAGCCATTAA ATTTTAAAA ACTTTTAAAA			929
	ACGATAAACA TAATCCGCGC TCCAAGTAAC ATAGCTTTCA AAAATG			975

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:

- 35 (A) LENGTH: 201 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

Met Ser Leu Leu Ala Thr Leu Leu Leu Ala Ser Cys Leu Pro Pro Lys
5 1 5 10 15
Gly His His Ser Gly Leu Val Asn Leu Tyr Ile Ala His Gln Gly Gln
20 25 30
Ser Val Arg Thr Tyr Trp Arg Lys Val Asp Arg Gly Val Ile Ala Lys
35 40 45
10 His Asn Glu Ala Leu Lys Lys Asp Pro Lys Ala Lys Leu Lys Asp Pro
50 55 60
Arg Gly Pro Leu Phe Met Leu Gly Ser Glu Arg Phe Met Leu Leu Trp
65 70 75 80
Lys Asn Arg Tyr Ala Leu Ala Lys Pro Gln Ser Phe Arg Leu Glu Pro
15 85 90 95
Gly Phe Tyr Tyr Leu Asp Ser Phe Ser Val Glu Thr Gln Lys Gly Val
100 105 110
Leu Gln Ser Ala Pro Gly Tyr Ser Tyr Thr Lys Asn Gly Tyr Asp Phe
115 120 125
20 Lys Asn Asn Arg Pro Phe Phe Leu Ala Phe Glu Val Lys Pro Asp Gly
130 135 140
Lys Thr Ile Leu Pro Ser Val Glu Leu Ser Leu Ile Lys Thr Pro Arg
145 150 155 160
Gly Phe Leu Gly Val Phe Leu Phe Asp Asn Asn Glu Lys Gly Thr Asn
25 165 170 175
Ala Lys Trp Ile Glu Gly Ser Leu Asn Leu Lys Leu Lys Asn Ala Ser
180 185 190
Phe Lys Asp Ala Trp Gly Leu Glu Gln
195 200

30 (2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1116 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
(B) LOCATION: 90...1076
(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

TATAAATACA TCGTTTCATT AGCGAATTTA ATGGCGTTAA GCGATCATAT TGATTTATTT

60

	TATGAATTTG TTTATTAAGG GAAAAAATC ATG TCA AAT AGC ATG TTG GAT AAA	113
	Met Ser Asn Ser Met Leu Asp Lys	
	1 5	
5	AAT AAA GCG ATT CTT ACA GGG GGT GGG GCT TTA TTA TTA GGG CTA ATC	161
	Asn Lys Ala Ile Leu Thr Gly Gly Gly Ala Leu Leu Leu Gly Leu Ile	
	10 15 20	
	GTG CTT TTT TAT TTA GCT TAT CGC CCT AAG GCT GAA GTG TTG CAA GGG	209
	Val Leu Phe Tyr Leu Ala Tyr Arg Pro Lys Ala Glu Val Leu Gln Gly	
	25 30 35 40	
10	TTT TTG GAA GCC AGA GAA TAC AGC GTG AGC TCC AAA GTC CCT GGC CGC	257
	Phe Leu Glu Ala Arg Glu Tyr Ser Val Ser Ser Lys Val Pro Gly Arg	
	45 50 55	
	ATT GAA AAG GTG TTT GTT AAA AAA GGC GAT CAC ATT AAA AAG GGC GAT	305
	Ile Glu Lys Val Phe Val Lys Lys Gly Asp His Ile Lys Lys Gly Asp	
15	60 65 70	
	TTG GTT TTT AGC ATT TCT AGC CCT GAA TTA GAA GCC AAA CTC GCT CAA	353
	Leu Val Phe Ser Ile Ser Ser Pro Glu Leu Glu Ala Lys Leu Ala Gln	
	75 80 85	
	GCT GAA GCC GGG CAT AAA GCC GCT AAA GCG CTT AGC GAT GAA GTC AAA	401
20	Ala Glu Ala Gly His Lys Ala Ala Lys Ala Leu Ser Asp Glu Val Lys	
	90 95 100	
	AGA GGC TCA AGA GAC GAA ACG ATT AAT TCT GCG AGA GAC GTT TGG CAA	449
	Arg Gly Ser Arg Asp Glu Thr Ile Asn Ser Ala Arg Asp Val Trp Gln	
	105 110 115 120	
25	GCA GCC AAA TCC CAA GCC ACT TTA GCC AAA GAG ACT TAT AAG CGC GTT	497
	Ala Ala Lys Ser Gln Ala Thr Leu Ala Lys Glu Thr Tyr Lys Arg Val	
	125 130 135	
	CAA GAT TTG TAT GAT AAT GGC GTG GCG AGC TTG CAA AAG CGC GAT GAA	545
30	Gln Asp Leu Tyr Asp Asn Gly Val Ala Ser Leu Gln Lys Arg Asp Glu	
	140 145 150	
	GCC TAT GCG GCT TAT GAA AGC ACT AAA TAC AAC GAG AGC GCG GCT TAC	593
	Ala Tyr Ala Ala Tyr Glu Ser Thr Lys Tyr Asn Glu Ser Ala Ala Tyr	
	155 160 165	
	CAA AAG TAT AAA ATG GCT TTA GGG GGG GCG AGC TCT GAA AGT AAG ATT	641
35	Gln Lys Tyr Lys Met Ala Leu Gly Gly Ala Ser Ser Glu Ser Lys Ile	
	170 175 180	
	GCC GCT AAG GCT AAA GAG AGC GCG GCT TTA GGG CAA GTG AAT GAA GTG	689
	Ala Ala Lys Ala Lys Glu Ser Ala Ala Leu Gly Gln Val Asn Glu Val	

	185	190	195	200	
	GAG TCT TAT TTA AAA GAC GTC AAA GCG ACA GCC CCA ATT GAT GGG GAA				737
	Glu Ser Tyr Leu Lys Asp Val Lys Ala Thr Ala Pro Ile Asp Gly Glu	205	210	215	
5	GTG AGT AAC GTG CTT TTA AGC GGT GGC GAG CTT AGC CCT AAG GGT TTT				785
	Val Ser Asn Val Leu Leu Ser Gly Gly Glu Leu Ser Pro Lys Gly Phe	220	225	230	
	CCT GTG GTT TTA ATG ATA GAT TTA AAG GAT AGT TGG TTA AAA ATC AGC				833
10	Pro Val Val Leu Met Ile Asp Leu Lys Asp Ser Trp Leu Lys Ile Ser	235	240	245	
	GTG CCT GAA AAG TAT TTG AAC GAG TTT AAA GTG GGT AAG GAA TTT GAA				881
	Val Pro Glu Lys Tyr Leu Asn Glu Phe Lys Val Gly Lys Glu Phe Glu	250	255	260	
	GGC TAT ATC CCG GCG TTG AAA AAA AGC ACG AAA TTC AGG GTC AAA TAT				929
15	Gly Tyr Ile Pro Ala Leu Lys Lys Ser Thr Lys Phe Arg Val Lys Tyr	265	270	275	280
	TTG AGC GTG ATG GGG GAT TTT GCG ACT TGG AAA GCG ACG AAT AAT TCC				977
	Leu Ser Val Met Gly Asp Phe Ala Thr Trp Lys Ala Thr Asn Asn Ser	285	290	295	
20	AAC ACT TAC GAC ATG AAA AGC TAT GAA GTG GAA GCC ATA CCC TTA GAA				1025
	Asn Thr Tyr Asp Met Lys Ser Tyr Glu Val Glu Ala Ile Pro Leu Glu	300	305	310	
	GAG TTG GAA AAT TTT AGG GTA GGG ATG AGC GTG TTA GTT ACC ATT AAA				1073
25	Glu Leu Glu Asn Phe Arg Val Gly Met Ser Val Leu Val Thr Ile Lys	315	320	325	
	CCT TAAAAAGGAT TGTTTTGTTC AGATTGATAA GCGCATGGGT				1116
	Pro				

(2) INFORMATION FOR SEQ ID NO:52:

- 30 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 329 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- 35 (ii) MOLECULE TYPE: protein
- (v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

	Met	Ser	Asn	Ser	Met	Leu	Asp	Lys	Asn	Lys	Ala	Ile	Leu	Thr	Gly	Gly	
	1				5					10					15		
5	Gly	Ala	Leu	Leu	Leu	Gly	Leu	Ile	Val	Leu	Phe	Tyr	Leu	Ala	Tyr	Arg	
				20					25					30			
	Pro	Lys	Ala	Glu	Val	Leu	Gln	Gly	Phe	Leu	Glu	Ala	Arg	Glu	Tyr	Ser	
			35					40					45				
	Val	Ser	Ser	Lys	Val	Pro	Gly	Arg	Ile	Glu	Lys	Val	Phe	Val	Lys	Lys	
		50					55					60					
10	Gly	Asp	His	Ile	Lys	Lys	Gly	Asp	Leu	Val	Phe	Ser	Ile	Ser	Ser	Pro	
	65					70					75					80	
	Glu	Leu	Glu	Ala	Lys	Leu	Ala	Gln	Ala	Glu	Ala	Gly	His	Lys	Ala	Ala	
					85					90					95		
	Lys	Ala	Leu	Ser	Asp	Glu	Val	Lys	Arg	Gly	Ser	Arg	Asp	Glu	Thr	Ile	
15				100					105				110				
	Asn	Ser	Ala	Arg	Asp	Val	Trp	Gln	Ala	Ala	Lys	Ser	Gln	Ala	Thr	Leu	
			115					120					125				
	Ala	Lys	Glu	Thr	Tyr	Lys	Arg	Val	Gln	Asp	Leu	Tyr	Asp	Asn	Gly	Val	
		130					135					140					
20	Ala	Ser	Leu	Gln	Lys	Arg	Asp	Glu	Ala	Tyr	Ala	Ala	Tyr	Glu	Ser	Thr	
	145					150					155					160	
	Lys	Tyr	Asn	Glu	Ser	Ala	Ala	Tyr	Gln	Lys	Tyr	Lys	Met	Ala	Leu	Gly	
				165					170				175				
	Gly	Ala	Ser	Ser	Glu	Ser	Lys	Ile	Ala	Ala	Lys	Ala	Lys	Glu	Ser	Ala	
25				180					185				190				
	Ala	Leu	Gly	Gln	Val	Asn	Glu	Val	Glu	Ser	Tyr	Leu	Lys	Asp	Val	Lys	
		195						200				205					
	Ala	Thr	Ala	Pro	Ile	Asp	Gly	Glu	Val	Ser	Asn	Val	Leu	Leu	Ser	Gly	
		210					215					220					
30	Gly	Glu	Leu	Ser	Pro	Lys	Gly	Phe	Pro	Val	Val	Leu	Met	Ile	Asp	Leu	
	225					230					235					240	
	Lys	Asp	Ser	Trp	Leu	Lys	Ile	Ser	Val	Pro	Glu	Lys	Tyr	Leu	Asn	Glu	
				245						250					255		
	Phe	Lys	Val	Gly	Lys	Glu	Phe	Glu	Gly	Tyr	Ile	Pro	Ala	Leu	Lys	Lys	
35				260					265				270				
	Ser	Thr	Lys	Phe	Arg	Val	Lys	Tyr	Leu	Ser	Val	Met	Gly	Asp	Phe	Ala	
			275					280					285				
	Thr	Trp	Lys	Ala	Thr	Asn	Asn	Ser	Asn	Thr	Tyr	Asp	Met	Lys	Ser	Tyr	
		290					295					300					
40	Glu	Val	Glu	Ala	Ile	Pro	Leu	Glu	Glu	Leu	Glu	Asn	Phe	Arg	Val	Gly	
	305					310					315					320	
	Met	Ser	Val	Leu	Val	Thr	Ile	Lys	Pro								
						325											

(2) INFORMATION FOR SEQ ID NO:53:

- 45 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1514 base pairs
 (B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence

5 (B) LOCATION: 94...1467

(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

	AAATAAAATA GCGATCATT TAACATGTTG CTTTTTAAGT GAAAGCGTTA AGTTGTTAGG	60
10	GTATAGTGGC TTAAAAATTT TAGGATATTG AGA ATG CTT GAA ACT TCT AGC CAT	114
	Met Leu Glu Thr Ser Ser His	
	1 5	
	TTT TTA AAA TCG TTT CGC TTG AAG CGT TAT ATA GGG TTT TTA TTG ATT	162
	Phe Leu Lys Ser Phe Arg Leu Lys Arg Tyr Ile Gly Phe Leu Leu Ile	
15	10 15 20	
	TCT TTA GCG TTA TTA ATC ACG CCC TTT GTT CGC ATT GAT GGG GCG CAT	210
	Ser Leu Ala Leu Leu Ile Thr Pro Phe Val Arg Ile Asp Gly Ala His	
	25 30 35	
	TTG TTT TTG ATC TCT TTT GAG CAT AAG CAA CTG CAT TTT TTA GGC AAG	258
20	Leu Phe Leu Ile Ser Phe Glu His Lys Gln Leu His Phe Leu Gly Lys	
	40 45 50 55	
	ATC TTT AGC GCT GAA GAA TTG CAA GTC ATG CCT TTT ATG GTT ATT TTG	306
	Ile Phe Ser Ala Glu Glu Leu Gln Val Met Pro Phe Met Val Ile Leu	
	60 65 70	
25	CTT TTT ATA GGG ATT TTT TTC ATC ACC ACT AGC CTT GGG CGT GTG TGG	354
	Leu Phe Ile Gly Ile Phe Phe Ile Thr Ser Leu Gly Arg Val Trp	
	75 80 85	
	TGC GGT TGG GCT TGC CCG CAA ACC TTT TTA AGG GTG CTT TAT AGA GAT	402
	Cys Gly Trp Ala Cys Pro Gln Thr Phe Leu Arg Val Leu Tyr Arg Asp	
30	90 95 100	
	GTG ATT GAA ACC AAG ATT TTC AAA CTC CAT AAA AAG ATC AGC AAC AAG	450
	Val Ile Glu Thr Lys Ile Phe Lys Leu His Lys Lys Ile Ser Asn Lys	
	105 110 115	
	CAA GAA AGC CCT AAA AAC ACC CCA AGC TAC AAG ATC CGT AAA GTA TTG	498
35	Gln Glu Ser Pro Lys Asn Thr Pro Ser Tyr Lys Ile Arg Lys Val Leu	
	120 125 130 135	
	AGC GTT TTA TTG TTC GCT CCT GTT GTG GCG GGG CTA ATG ATG TTG TTT	546

Ser Val Leu Leu Phe Ala Pro Val Val Ala Gly Leu Met Met Leu Phe
140 145 150

TTC TTT TAT TTC ATC GCC CCA GAA GAT TTT TTT ATG TAT CTT AAA AAC 594
Phe Phe Tyr Phe Ile Ala Pro Glu Asp Phe Phe Met Tyr Leu Lys Asn
5 155 160 165

CCT AGC GAT CAC CCT ATT GCT ATG GGT TTT TGG CTT TTT AGC ACG GCT 642
Pro Ser Asp His Pro Ile Ala Met Gly Phe Trp Leu Phe Ser Thr Ala
170 175 180

GTG GTG CTA TTT GAT ATA GTG GTG GTT GCG GAG CGT TTT TGC ATT TAT 690
10 Val Val Leu Phe Asp Ile Val Val Val Ala Glu Arg Phe Cys Ile Tyr
185 190 195

TTA TGC CCT TAC GCT AGG GTG CAA TCG GTG TTG TAT GAC AAT GAC ACC 738
Leu Cys Pro Tyr Ala Arg Val Gln Ser Val Leu Tyr Asp Asn Asp Thr
200 205 210 215

TTA AAC CCT ATT TAT GAT GAA AAG CGC GGC GGA GCG CTT TAT AAT AAT 786
15 Leu Asn Pro Ile Tyr Asp Glu Lys Arg Gly Gly Ala Leu Tyr Asn Asn
220 225 230

CAG GGC CAT CTC TTC CCC TTA CCT CCC AAA AAA CGC AGC CCA GAA AAC 834
20 Gln Gly His Leu Phe Pro Leu Pro Pro Lys Lys Arg Ser Pro Glu Asn
235 240 245

GAA TGC GTG AAT TGT TTG CAT TGC GTG CAG GTT TGC CCC ACG CAT ATT 882
Glu Cys Val Asn Cys Leu His Cys Val Gln Val Cys Pro Thr His Ile
250 255 260

GAC ATC AGG AAG GGC TTG CAA TTA GAA TGC ATC AAT TGT TTA GAA TGC 930
25 Asp Ile Arg Lys Gly Leu Gln Leu Glu Cys Ile Asn Cys Leu Glu Cys
265 270 275

GTG GAT GCA TGC ACG ATT ACC ATG GCT AAA TTT AAC CGC CCT TCA CTC 978
Val Asp Ala Cys Thr Ile Thr Met Ala Lys Phe Asn Arg Pro Ser Leu
280 285 290 295

ATC CAA TGG TCT TCA ACT AAC GCT ATT AAT ACG CGC CAA AAA GTG CAC 1026
30 Ile Gln Trp Ser Ser Thr Asn Ala Ile Asn Thr Arg Gln Lys Val His
300 305 310

CTG GTG CGT TTA AAA ACG ATC GCT TAC ATG GGG GTT ATC GCT ATT GTG 1074
35 Leu Val Arg Leu Lys Thr Ile Ala Tyr Met Gly Val Ile Ala Ile Val
315 320 325

ATC GCT CTT TTA GCC ATC ACT TCG TTT AAA AAA GAA CGC ATG CTC TTA 1122
Ile Ala Leu Leu Ala Ile Thr Ser Phe Lys Lys Glu Arg Met Leu Leu
330 335 340

	GAC ATT AAC CGC AAC AGC GAT CTG TAT GAA TTG CGC TCT AGC GGG TAT	1170
	Asp Ile Asn Arg Asn Ser Asp Leu Tyr Glu Leu Arg Ser Ser Gly Tyr	
	345 350 355	
5	GTG GAT AAC GAT TAC GTG TTT TTA TTC CAC AAC ACG GAC AAT AAA GAC	1218
	Val Asp Asn Asp Tyr Val Phe Leu Phe His Asn Thr Asp Asn Lys Asp	
	360 365 370 375	
	CAT GAG TTT TAT TTC AAA GTT TTA GGG CAA AAA GAC ATT CAG ATC AAA	1266
	His Glu Phe Tyr Phe Lys Val Leu Gly Gln Lys Asp Ile Gln Ile Lys	
	380 385 390	
10	AAG CCT TTA AAT CCT ATC GCC ATT AAA GCC GGG CAA AAG ATT AAA GCG	1314
	Lys Pro Leu Asn Pro Ile Ala Ile Lys Ala Gly Gln Lys Ile Lys Ala	
	395 400 405	
15	GTA GTG ATT TTA AGA AAA CCC CTA AAG AGT AAC GCC ACA GAA TAC AAG	1362
	Val Val Ile Leu Arg Lys Pro Leu Lys Ser Asn Ala Thr Glu Tyr Lys	
	410 415 420	
	AAC GCT AAA GAC GCT CTA ATC CCC ATT ACC ATA CAA GCT TAT AGC GCG	1410
	Asn Ala Lys Asp Ala Leu Ile Pro Ile Thr Ile Gln Ala Tyr Ser Ala	
	425 430 435	
20	GAC GAT AAG AAT ATT ACG ATA GAA AGG GAA TCG GTG TTT ATT GCA CCA	1458
	Asp Asp Lys Asn Ile Thr Ile Glu Arg Glu Ser Val Phe Ile Ala Pro	
	440 445 450 455	
	AGT GAG GAT TGAAGCCTAA AACTAGCGTT CAATCACTTC ATAAGGCAAG CCTTGTT	1514
	Ser Glu Asp	

(2) INFORMATION FOR SEQ ID NO:54:

25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 458 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

30 (ii) MOLECULE TYPE: protein
 (v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

	Met Leu Glu Thr Ser Ser His Phe Leu Lys Ser Phe Arg Leu Lys Arg
	1 5 10 15
35	Tyr Ile Gly Phe Leu Leu Ile Ser Leu Ala Leu Leu Ile Thr Pro Phe
	20 25 30

	Val	Arg	Ile	Asp	Gly	Ala	His	Leu	Phe	Leu	Ile	Ser	Phe	Glu	His	Lys
		35						40					45			
	Gln	Leu	His	Phe	Leu	Gly	Lys	Ile	Phe	Ser	Ala	Glu	Glu	Leu	Gln	Val
		50					55					60				
5	Met	Pro	Phe	Met	Val	Ile	Leu	Leu	Phe	Ile	Gly	Ile	Phe	Phe	Ile	Thr
	65					70				75					80	
	Thr	Ser	Leu	Gly	Arg	Val	Trp	Cys	Gly	Trp	Ala	Cys	Pro	Gln	Thr	Phe
					85					90					95	
	Leu	Arg	Val	Leu	Tyr	Arg	Asp	Val	Ile	Glu	Thr	Lys	Ile	Phe	Lys	Leu
10					100				105					110		
	His	Lys	Lys	Ile	Ser	Asn	Lys	Gln	Glu	Ser	Pro	Lys	Asn	Thr	Pro	Ser
			115					120					125			
	Tyr	Lys	Ile	Arg	Lys	Val	Leu	Ser	Val	Leu	Leu	Phe	Ala	Pro	Val	Val
		130					135					140				
15	Ala	Gly	Leu	Met	Met	Leu	Phe	Phe	Phe	Tyr	Phe	Ile	Ala	Pro	Glu	Asp
	145					150					155				160	
	Phe	Phe	Met	Tyr	Leu	Lys	Asn	Pro	Ser	Asp	His	Pro	Ile	Ala	Met	Gly
					165					170					175	
	Phe	Trp	Leu	Phe	Ser	Thr	Ala	Val	Val	Leu	Phe	Asp	Ile	Val	Val	Val
20					180					185				190		
	Ala	Glu	Arg	Phe	Cys	Ile	Tyr	Leu	Cys	Pro	Tyr	Ala	Arg	Val	Gln	Ser
			195					200					205			
	Val	Leu	Tyr	Asp	Asn	Asp	Thr	Leu	Asn	Pro	Ile	Tyr	Asp	Glu	Lys	Arg
		210					215					220				
25	Gly	Gly	Ala	Leu	Tyr	Asn	Asn	Gln	Gly	His	Leu	Phe	Pro	Leu	Pro	Pro
	225					230					235				240	
	Lys	Lys	Arg	Ser	Pro	Glu	Asn	Glu	Cys	Val	Asn	Cys	Leu	His	Cys	Val
					245					250					255	
	Gln	Val	Cys	Pro	Thr	His	Ile	Asp	Ile	Arg	Lys	Gly	Leu	Gln	Leu	Glu
30					260					265				270		
	Cys	Ile	Asn	Cys	Leu	Glu	Cys	Val	Asp	Ala	Cys	Thr	Ile	Thr	Met	Ala
			275					280					285			
	Lys	Phe	Asn	Arg	Pro	Ser	Leu	Ile	Gln	Trp	Ser	Ser	Thr	Asn	Ala	Ile
		290					295					300				
35	Asn	Thr	Arg	Gln	Lys	Val	His	Leu	Val	Arg	Leu	Lys	Thr	Ile	Ala	Tyr
	305					310					315				320	
	Met	Gly	Val	Ile	Ala	Ile	Val	Ile	Ala	Leu	Leu	Ala	Ile	Thr	Ser	Phe
					325					330					335	
	Lys	Lys	Glu	Arg	Met	Leu	Leu	Asp	Ile	Asn	Arg	Asn	Ser	Asp	Leu	Tyr
40					340					345				350		
	Glu	Leu	Arg	Ser	Ser	Gly	Tyr	Val	Asp	Asn	Asp	Tyr	Val	Phe	Leu	Phe
			355					360					365			
	His	Asn	Thr	Asp	Asn	Lys	Asp	His	Glu	Phe	Tyr	Phe	Lys	Val	Leu	Gly
		370					375					380				
45	Gln	Lys	Asp	Ile	Gln	Ile	Lys	Lys	Pro	Leu	Asn	Pro	Ile	Ala	Ile	Lys
	385					390					395				400	
	Ala	Gly	Gln	Lys	Ile	Lys	Ala	Val	Val	Ile	Leu	Arg	Lys	Pro	Leu	Lys
					405					410					415	
	Ser	Asn	Ala	Thr	Glu	Tyr	Lys	Asn	Ala	Lys	Asp	Ala	Leu	Ile	Pro	Ile
50					420					425					430	

Thr Ile Gln Ala Tyr Ser Ala Asp Asp Lys Asn Ile Thr Ile Glu Arg
 435 440 445
 Glu Ser Val Phe Ile Ala Pro Ser Glu Asp
 450 455

5 (2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 990 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 228...782
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

	ACGATTTGAT	CAATAACGAA	AATAAAATTG	ATGAAATCAA	TAATGAAGAA	AACGCTGATC	60
	CTTCGCAAAA	AAGAACGAAC	AACGTTTTC	AACGAGCCAC	TAACCACCAA	GACAATCTCA	120
	ATTCCCCACT	CAACAGGAAG	TATTAAAGTG	TGAAACTTTT	TTCAAAGGAT	TTATTTAAAA	180
20	AAGTAACCCC	TTTATTTTTA	AGCGTTTATT	TTTAAACCC	CACCATT	ATG CAA GCC	236
					Met Gln Ala		
					1		
	AAA AGC CGT	TTT TAT GTG	GCT TCT CAA	TAC CAG GTG	GGG AAA ATG	ATC	284
	Lys Ser Arg	Phe Tyr Val	Ala Ser Gln	Tyr Gln Val	Gly Lys Met	Ile	
25	5	10	15				
	ATG AAA AAA	TAC AAC GAT	CTC AAA CGC	ACG ATT GAA	GGG GCG AGC	TTT	332
	Met Lys Lys	Tyr Asn Asp	Leu Lys Arg	Thr Ile Glu	Gly Ala Ser	Phe	
	20	25	30	35			
	TCT TTA GGC	TGG GAG ATT	AAC CCC ACT	AAC TAC TGG	TTT TAT TCG	CGC	380
30	Ser Leu Gly	Trp Glu Ile	Asn Pro Thr	Asn Tyr Trp	Phe Tyr Ser	Arg	
	40	45	50				
	TAT TAC TTT	TTT ATG GAT	TAC GGG AAT	GTC ATT CTC	AAT AAA AGA	ACG	428
	Tyr Tyr Phe	Phe Met Asp	Tyr Gly Asn	Val Ile Leu	Asn Lys Arg	Thr	
	55	60	65				
35	GGC GCT CAA	GCG AAC ATG	TTC ACT TAT	GGC TTT GGG	GGG GAT TTG	ATT	476
	Gly Ala Gln	Ala Asn Met	Phe Thr Tyr	Gly Phe Gly	Gly Asp Leu	Ile	
	70	75	80				
	GTG GAA TAC	AAT AAA AAC	CCC TTG TAT	GTA TTT TCT	CTT TTT TAT	GGC	524

Val Glu Tyr Asn Lys Asn Pro Leu Tyr Val Phe Ser Leu Phe Tyr Gly
85 90 95

ATG CAA GTT GCT GAA AAC ACA TGG ACG ATT TCC AAA CAC AGC GCG AAT 572
Met Gln Val Ala Glu Asn Thr Trp Thr Ile Ser Lys His Ser Ala Asn
5 100 105 110 115

TTC ATC ATT GAC GAT TGG CGC AGC ATT CAA GGG TTT TCG CTC AAA ACT 620
Phe Ile Ile Asp Asp Trp Arg Ser Ile Gln Gly Phe Ser Leu Lys Thr
120 125 130

TCC AAT TTT AGG ATG TTG GGT TTA GTG GGG TTT AAA TTC CAA ACC GTG 668
10 Ser Asn Phe Arg Met Leu Gly Leu Val Gly Phe Lys Phe Gln Thr Val
135 140 145

CTA TTC CAC CAT GAC GCA AGT ATT GAA GTG GGG ATC AAA TGG CCT TTT 716
Leu Phe His His Asp Ala Ser Ile Glu Val Gly Ile Lys Trp Pro Phe
150 155 160

15 GCT TTT GAA TAC GAC TCA GCC TTT GTA AGG CTT TTT TCT GTC TTT ATT 764
Ala Phe Glu Tyr Asp Ser Ala Phe Val Arg Leu Phe Ser Val Phe Ile
165 170 175

TCG CAC ACT TTC TAC CTT TAACTAATT CCAACCCTAC CGGGCAATGA TCGCTCCC 820
20 Ser His Thr Phe Tyr Leu
180 185

TAAATATCT TTATAGATTA AAGCGTCTTT TAAGCGCGTT TTAAAGGGT TAGAGCATAA 880
AAAATAATCA ATGCGCCAAC CAATGTTTTT ATCCCTTGCT TGTTGCATGT AACTCCACCA 940
GGTGTAAAGCC TTTTCTTTGT TAGGGTAAAA ATAACGGAAA GTGTCAATAA 990

(2) INFORMATION FOR SEQ ID NO:56:

25 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 185 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

30 (ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

Met Gln Ala Lys Ser Arg Phe Tyr Val Ala Ser Gln Tyr Gln Val Gly
1 5 10 15
35 Lys Met Ile Met Lys Lys Tyr Asn Asp Leu Lys Arg Thr Ile Glu Gly
20 25 30
Ala Ser Phe Ser Leu Gly Trp Glu Ile Asn Pro Thr Asn Tyr Trp Phe
35 40 45

	Tyr	Ser	Arg	Tyr	Tyr	Phe	Phe	Met	Asp	Tyr	Gly	Asn	Val	Ile	Leu	Asn
	50						55					60				
	Lys	Arg	Thr	Gly	Ala	Gln	Ala	Asn	Met	Phe	Thr	Tyr	Gly	Phe	Gly	Gly
	65					70					75				80	
5	Asp	Leu	Ile	Val	Glu	Tyr	Asn	Lys	Asn	Pro	Leu	Tyr	Val	Phe	Ser	Leu
				85					90					95		
	Phe	Tyr	Gly	Met	Gln	Val	Ala	Glu	Asn	Thr	Trp	Thr	Ile	Ser	Lys	His
				100					105					110		
	Ser	Ala	Asn	Phe	Ile	Ile	Asp	Asp	Trp	Arg	Ser	Ile	Gln	Gly	Phe	Ser
10			115					120					125			
	Leu	Lys	Thr	Ser	Asn	Phe	Arg	Met	Leu	Gly	Leu	Val	Gly	Phe	Lys	Phe
		130					135					140				
	Gln	Thr	Val	Leu	Phe	His	His	Asp	Ala	Ser	Ile	Glu	Val	Gly	Ile	Lys
	145					150					155				160	
15	Trp	Pro	Phe	Ala	Phe	Glu	Tyr	Asp	Ser	Ala	Phe	Val	Arg	Leu	Phe	Ser
					165					170					175	
	Val	Phe	Ile	Ser	His	Thr	Phe	Tyr	Leu							
				180					185							

(2) INFORMATION FOR SEQ ID NO:57:

20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1161 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

25 (ix) FEATURE:
 (A) NAME/KEY: Coding Sequence
 (B) LOCATION: 109...1113
 (D) OTHER INFORMATION:

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

	ATCTTACCTT	TATCTTTTAA	GATTTTATGA	AAAATAGTTT	CATTTTTTACT	ATTGTTATTT	60
	TCTTAGTAAT	GTTATAATCG	CTTTATAAAT	CATACAAAAA	GGATCGCT	ATG TTA GTT	117
						Met Leu Val	
						1	
35	ACT CGC TTT AAA AAA	GCT TTC ATT TCT	TAT TCT TTA	GGC GTG CTT	GTC		165
	Thr Arg Phe Lys Lys	Ala Phe Ile Ser	Tyr Ser Leu	Gly Val Leu	Val		
	5	10	15				
	GCT TCA TTA TGG	TTG AAC GTG	TGC AAC GCT	TCA GCG CAA	GAA GTC AAA		213
	Ala Ser Leu Trp	Leu Asn Val	Cys Asn Ala	Ser Ala Gln	Glu Val Lys		
40	20	25	30	35			
	GTC AAG GAT	TAT TTC GGG	GAG CAA ACC	ATC AAG CTT	CCT GTT TCT	AAA	261

	Val Lys Asp Tyr Phe Gly Glu Gln Thr Ile Lys Leu Pro Val Ser Lys	
	40 45 50	
5	ATA GCC TAT ATA GGG AGC TAT GTA GAA GTG CCT GCC ATG CTT AAT GTT Ile Ala Tyr Ile Gly Ser Tyr Val Glu Val Pro Ala Met Leu Asn Val	309
	55 60 65	
	TGG AAT AGG GTT GTA GGC GTT TCG GAT TAC GCT TTT AAA GAC GAT ATT Trp Asn Arg Val Val Gly Val Ser Asp Tyr Ala Phe Lys Asp Asp Ile	357
	70 75 80	
10	GTC AAA GCC ACT CTC AAA GGC GAA GAT CTT AAA CGC GTC AAA CAC ATG Val Lys Ala Thr Leu Lys Gly Glu Asp Leu Lys Arg Val Lys His Met	405
	85 90 95	
	AGC ACT GAT CAT ACA GCC GCG CTA AAT GTA GAG CTT TTA AAA AAG CTT Ser Thr Asp His Thr Ala Ala Leu Asn Val Glu Leu Leu Lys Lys Leu	453
	100 105 110 115	
15	AGC CCT GAT CTT GTG GTA ACC TTT GTG GGC AAC CCT AAA GCG GTA GAG Ser Pro Asp Leu Val Val Thr Phe Val Gly Asn Pro Lys Ala Val Glu	501
	120 125 130	
20	CAT GCG AAA AAA TTT GGT ATA TCA TTT CTT TCT TTT CAA GAG ACA ACG His Ala Lys Lys Phe Gly Ile Ser Phe Leu Ser Phe Gln Glu Thr Thr	549
	135 140 145	
	ATT GCA GAG GCC ATG CAG GCC ATG CAA GCT CAA GCC ACG GTT TTA GAG Ile Ala Glu Ala Met Gln Ala Met Gln Ala Gln Ala Thr Val Leu Glu	597
	150 155 160	
25	ATT GAC GCT TCC AAA AAA TTC GCC AAA ATG CAA GAA ACT TTG GAT TTT Ile Asp Ala Ser Lys Lys Phe Ala Lys Met Gln Glu Thr Leu Asp Phe	645
	165 170 175	
	ATT GCT GAG CGT TTG AAA AAT GTC AAA AAG AAA AAG GGG GTG GAG CTT Ile Ala Glu Arg Leu Lys Asn Val Lys Lys Lys Lys Gly Val Glu Leu	693
	180 185 190 195	
30	TTC CAT AAA GCC AAT AAA ATC AGC GGC CAT CAA GCC ATT AGC TCA GAC Phe His Lys Ala Asn Lys Ile Ser Gly His Gln Ala Ile Ser Ser Asp	741
	200 205 210	
35	ATT TTA GAA AAA GGG GGC ATA GAC AAT TTT GGC TTG AAA TAT GTC AAA Ile Leu Glu Lys Gly Gly Ile Asp Asn Phe Gly Leu Lys Tyr Val Lys	789
	215 220 225	
	TTT GGG CGT GCT GAC ATT AGC GTG GAA AAA ATC GTT AAA GAA AAC CCT Phe Gly Arg Ala Asp Ile Ser Val Glu Lys Ile Val Lys Glu Asn Pro	837
	230 235 240	

GAG ATT ATC TTT ATT TGG TGG ATA AGC CCA CTC ACG CCT GAA GAT GTG 885
 Glu Ile Ile Phe Ile Trp Trp Ile Ser Pro Leu Thr Pro Glu Asp Val
 245 250 255

TTA AAC AAC CCC AAA TTT GCT ACC ATC AAA GCC ATT AAA AAC AAG CAG 933
 5 Leu Asn Asn Pro Lys Phe Ala Thr Ile Lys Ala Ile Lys Asn Lys Gln
 260 265 270 275

GTT TAT AAA CTC CCC ACA ATG GAT ATT GGC GGG CCT AGA GCC CCA CTC 981
 Val Tyr Lys Leu Pro Thr Met Asp Ile Gly Gly Pro Arg Ala Pro Leu
 280 285 290

10 ATA AGT CTT TTT ATC GCT CTA AAA GCC CAC CCT GAA GCC TTT AAG GGC 1029
 Ile Ser Leu Phe Ile Ala Leu Lys Ala His Pro Glu Ala Phe Lys Gly
 295 300 305

GTG GAT ATT AAT GCG ATG GTT AAA GAC TAC TAT AAA GTG GTT TTT GAT 1077
 Val Asp Ile Asn Ala Met Val Lys Asp Tyr Tyr Lys Val Val Phe Asp
 15 310 315 320

TTG AAT GAT GCA GAG GTT GAG CCC TTT TTA TGG CAT TAATTTTAA AAAGGG 1129
 Leu Asn Asp Ala Glu Val Glu Pro Phe Leu Trp His
 325 330 335

GTTGATGTTT TTAGCCTTTC GTGTATCGCG CT 1161

20 (2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 335 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

25 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

Met Leu Val Thr Arg Phe Lys Lys Ala Phe Ile Ser Tyr Ser Leu Gly
 30 1 5 10 15
 Val Leu Val Ala Ser Leu Trp Leu Asn Val Cys Asn Ala Ser Ala Gln
 20 25 30
 Glu Val Lys Val Lys Asp Tyr Phe Gly Glu Gln Thr Ile Lys Leu Pro
 35 40 45
 35 Val Ser Lys Ile Ala Tyr Ile Gly Ser Tyr Val Glu Val Pro Ala Met
 50 55 60
 Leu Asn Val Trp Asn Arg Val Val Gly Val Ser Asp Tyr Ala Phe Lys
 65 70 75 80
 Asp Asp Ile Val Lys Ala Thr Leu Lys Gly Glu Asp Leu Lys Arg Val

				85					90				95
	Lys	His	Met	Ser	Thr	Asp	His	Thr	Ala	Ala	Leu	Asn	Val
				100					105				110
5	Lys	Lys	Leu	Ser	Pro	Asp	Leu	Val	Val	Thr	Phe	Val	Gly
			115					120				125	
	Ala	Val	Glu	His	Ala	Lys	Lys	Phe	Gly	Ile	Ser	Phe	Leu
			130				135				140		
	Glu	Thr	Thr	Ile	Ala	Glu	Ala	Met	Gln	Ala	Met	Gln	Ala
						150				155			160
10	Val	Leu	Glu	Ile	Asp	Ala	Ser	Lys	Lys	Phe	Ala	Lys	Met
					165					170			175
	Leu	Asp	Phe	Ile	Ala	Glu	Arg	Leu	Lys	Asn	Val	Lys	Lys
				180				185				190	
	Val	Glu	Leu	Phe	His	Lys	Ala	Asn	Lys	Ile	Ser	Gly	His
15			195					200				205	
	Ser	Ser	Asp	Ile	Leu	Glu	Lys	Gly	Gly	Ile	Asp	Asn	Phe
			210				215				220		
	Tyr	Val	Lys	Phe	Gly	Arg	Ala	Asp	Ile	Ser	Val	Glu	Lys
					230					235			240
20	Glu	Asn	Pro	Glu	Ile	Ile	Phe	Ile	Trp	Trp	Ile	Ser	Pro
					245				250				255
	Glu	Asp	Val	Leu	Asn	Asn	Pro	Lys	Phe	Ala	Thr	Ile	Lys
				260				265				270	
	Asn	Lys	Gln	Val	Tyr	Lys	Leu	Pro	Thr	Met	Asp	Ile	Gly
25			275				280				285		
	Ala	Pro	Leu	Ile	Ser	Leu	Phe	Ile	Ala	Leu	Lys	Ala	His
			290				295				300		
	Phe	Lys	Gly	Val	Asp	Ile	Asn	Ala	Met	Val	Lys	Asp	Tyr
					310				315				320
30	Val	Phe	Asp	Leu	Asn	Asp	Ala	Glu	Val	Glu	Pro	Phe	Leu
				325				330				335	

(2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 800 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 121...669
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

TTATTTCGCAT GCATTAGCTA TTATTGAAGC TCAAAGCATT CAAGCGCATT TATTCTTAGA

60

	TGAAATCAAA CAAAGCCAAA AAGAAAAGAA AAAATTCCCC ACTTTCAAAG GAGGTTTTTA	120
	ATG CGT TGG TGG TGT TTT TTG GTG TGT TGT TTT GGT ATT TTA AGC GTG	168
	Met Arg Trp Trp Cys Phe Leu Val Cys Cys Phe Gly Ile Leu Ser Val	
	1 5 10 15	
5	ATG GAC GCT AAA AAA TTA GAG AAT AAG AAT TTG AAA AAA GAA AGA GAG	216
	Met Asp Ala Lys Lys Leu Glu Asn Lys Asn Leu Lys Lys Glu Arg Glu	
	20 25 30	
	CTT TTA GAG ATT ACT GGC AAC CAA TTT GTA GCG AAC GAC AAA ACC AAA	264
10	Leu Leu Glu Ile Thr Gly Asn Gln Phe Val Ala Asn Asp Lys Thr Lys	
	35 40 45	
	ACC GCT GTT ATT CAA GGC AAT GTG CAG ATC AAA AAG GGT AAA GAC CGG	312
	Thr Ala Val Ile Gln Gly Asn Val Gln Ile Lys Lys Gly Lys Asp Arg	
	50 55 60	
	TTG TTT GCG GAC AAG GTG AGC GTG TTT TTA AAC GAT AAA CGA AAG CCA	360
15	Leu Phe Ala Asp Lys Val Ser Val Phe Leu Asn Asp Lys Arg Lys Pro	
	65 70 75 80	
	GAG CGC TAT GAA GCC ACA GGG AAC ACG CAT TTT AAC ATC TTT ACA GAG	408
	Glu Arg Tyr Glu Ala Thr Gly Asn Thr His Phe Asn Ile Phe Thr Glu	
	85 90 95	
20	GAC AAT CGT GAA ATC AGC GGG AGT GCT GAC AAG CTC ATT TAT AAC GCG	456
	Asp Asn Arg Glu Ile Ser Gly Ser Ala Asp Lys Leu Ile Tyr Asn Ala	
	100 105 110	
	CTG AAT GGG GAA TAC AAA TTA TTG CAA AAT GCG GTG GTT AGA GAA GTG	504
25	Leu Asn Gly Glu Tyr Lys Leu Leu Gln Asn Ala Val Val Arg Glu Val	
	115 120 125	
	GGG AAA TCC AAT GTC ATC ACC GGC GAT GAA ATC ATT TTA AAC AAA ACT	552
	Gly Lys Ser Asn Val Ile Thr Gly Asp Glu Ile Ile Leu Asn Lys Thr	
	130 135 140	
	AAG GGT TAT GCT GAT GTG TTG GGG AGC GCG AAA CGG CCC GCT AAA TTC	600
30	Lys Gly Tyr Ala Asp Val Leu Gly Ser Ala Lys Arg Pro Ala Lys Phe	
	145 150 155 160	
	GTG TTT GAT ATG GAA GAT ATT AAT GAA GAA AAT CGT AAG GCT AAA TTG	648
	Val Phe Asp Met Glu Asp Ile Asn Glu Glu Asn Arg Lys Ala Lys Leu	
	165 170 175	
35	AAG AAG AAA GGC GAA AAA CCA TGATTGTCAT TAAAGACGCT CATTTTCTCA CTTC	703
	Lys Lys Lys Gly Glu Lys Pro	
	180	
	TTCAAGCCAA CTTTTTCAAT GCCCTGCGAG TTGACTTCT GAAATGGTGG TTTTAGGGCG	763

CAGCAATGTA GGCAAAAGCT CGTTTATTAA TACCTTG

800

(2) INFORMATION FOR SEQ ID NO:60:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 183 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

Met Arg Trp Trp Cys Phe Leu Val Cys Cys Phe Gly Ile Leu Ser Val
1 5 10 15
Met Asp Ala Lys Lys Leu Glu Asn Lys Asn Leu Lys Lys Glu Arg Glu
20 25 30
15 Leu Leu Glu Ile Thr Gly Asn Gln Phe Val Ala Asn Asp Lys Thr Lys
35 40 45
Thr Ala Val Ile Gln Gly Asn Val Gln Ile Lys Lys Gly Lys Asp Arg
50 55 60
Leu Phe Ala Asp Lys Val Ser Val Phe Leu Asn Asp Lys Arg Lys Pro
20 65 70 75 80
Glu Arg Tyr Glu Ala Thr Gly Asn Thr His Phe Asn Ile Phe Thr Glu
85 90 95
Asp Asn Arg Glu Ile Ser Gly Ser Ala Asp Lys Leu Ile Tyr Asn Ala
100 105 110
25 Leu Asn Gly Glu Tyr Lys Leu Leu Gln Asn Ala Val Val Arg Glu Val
115 120 125
Gly Lys Ser Asn Val Ile Thr Gly Asp Glu Ile Ile Leu Asn Lys Thr
130 135 140
Lys Gly Tyr Ala Asp Val Leu Gly Ser Ala Lys Arg Pro Ala Lys Phe
30 145 150 155 160
Val Phe Asp Met Glu Asp Ile Asn Glu Glu Asn Arg Lys Ala Lys Leu
165 170 175
Lys Lys Lys Gly Glu Lys Pro
180

(2) INFORMATION FOR SEQ ID NO:61:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 724 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence
 (B) LOCATION: 88...618
 (D) OTHER INFORMATION:

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

	GTGATGATTG AAAACATGTG AAAGAGCGTT TTTTAAGCTT TTAAATGGTG TTTGAATGCG	60
	AAAAAAGGC TAATACTATC ATAAGGA ATG AAG TTG ATA AAA TTT GTG CGT AAT	114
	Met Lys Leu Ile Lys Phe Val Arg Asn	
	1 5	
10	GTG GTT TTG TTC ATT TTA ACG GCG ATC TTT TTA GCG TTC ATG CTT TTG	162
	Val Val Leu Phe Ile Leu Thr Ala Ile Phe Leu Ala Phe Met Leu Leu	
	10 15 20 25	
	GTG AGT TAT TGC ATG CCC CAT TAT AGC GCG GCT GTC ATT AGC GGG GTG	210
	Val Ser Tyr Cys Met Pro His Tyr Ser Ala Ala Val Ile Ser Gly Val	
15	30 35 40	
	GAA GTC AAA AGA ATG AAT GAA AAT GAA AAC ACG CCC AAT AAT AAG GAA	258
	Glu Val Lys Arg Met Asn Glu Asn Glu Asn Thr Pro Asn Asn Lys Glu	
	45 50 55	
	GTA AAA ACC CTT GCT AGA GAT GTC TAT TTT GTG CAA ACT TAC GAC CCT	306
20	Val Lys Thr Leu Ala Arg Asp Val Tyr Phe Val Gln Thr Tyr Asp Pro	
	60 65 70	
	AAA GAT CAA AAA AGC GTA ACC GTT TAT CGT AAC GAA GAC ACG CGC TTT	354
	Lys Asp Gln Lys Ser Val Thr Val Tyr Arg Asn Glu Asp Thr Arg Phe	
	75 80 85	
25	AGC TTC CCT TTT TAT TTT AAG TTT AAT TCG GCT GAT ATT TCA GCC CTC	402
	Ser Phe Pro Phe Tyr Phe Lys Phe Asn Ser Ala Asp Ile Ser Ala Leu	
	90 95 100 105	
	GCT CAA AGT TTA ATC AAT CAG CAA GTG GAA GTG AAA TAC TAT GGT TGG	450
	Ala Gln Ser Leu Ile Asn Gln Gln Val Glu Val Lys Tyr Tyr Gly Trp	
30	110 115 120	
	CGG ATC AAT TTG TTT AAC ATG TTC CCT AAT GTG ATT TTT TTA AAG CCC	498
	Arg Ile Asn Leu Phe Asn Met Phe Pro Asn Val Ile Phe Leu Lys Pro	
	125 130 135	
	TTA AAA GAG AGC ACT GAC ATT TCA AAG CCC ATT TTT AGC TGG ATT TTA	546
35	Leu Lys Glu Ser Thr Asp Ile Ser Lys Pro Ile Phe Ser Trp Ile Leu	
	140 145 150	
	TAC GCT TTG CTG TTA ATG GGC TTT TTT ATC AGC GCG CGT TCT GTT TGC	594
	Tyr Ala Leu Leu Leu Met Gly Phe Phe Ile Ser Ala Arg Ser Val Cys	

155

160

165

ACT TTA TTT AAG AGC AAA GCT CAT TAAACTTTT AGGCTTTGTT GGAAAATCAC 648
 Thr Leu Phe Lys Ser Lys Ala His
 170 175

5 AATGGGGTTA TTGGAGCGTG TATTAAAAAG CTCAATATAG GGCAAGCTGA TGCTGTGAAA 708
 AGCGGTGTTG TTCCT 724

(2) INFORMATION FOR SEQ ID NO:62:

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 177 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

Met Lys Leu Ile Lys Phe Val Arg Asn Val Val Leu Phe Ile Leu Thr
 1 5 10 15
 Ala Ile Phe Leu Ala Phe Met Leu Leu Val Ser Tyr Cys Met Pro His
 20 25 30
 Tyr Ser Ala Ala Val Ile Ser Gly Val Glu Val Lys Arg Met Asn Glu
 35 40 45
 Asn Glu Asn Thr Pro Asn Asn Lys Glu Val Lys Thr Leu Ala Arg Asp
 50 55 60
 Val Tyr Phe Val Gln Thr Tyr Asp Pro Lys Asp Gln Lys Ser Val Thr
 25 65 70 75 80
 Val Tyr Arg Asn Glu Asp Thr Arg Phe Ser Phe Pro Phe Tyr Phe Lys
 85 90 95
 Phe Asn Ser Ala Asp Ile Ser Ala Leu Ala Gln Ser Leu Ile Asn Gln
 100 105 110
 30 Gln Val Glu Val Lys Tyr Tyr Gly Trp Arg Ile Asn Leu Phe Asn Met
 115 120 125
 Phe Pro Asn Val Ile Phe Leu Lys Pro Leu Lys Glu Ser Thr Asp Ile
 130 135 140
 Ser Lys Pro Ile Phe Ser Trp Ile Leu Tyr Ala Leu Leu Leu Met Gly
 35 145 150 155 160
 Phe Phe Ile Ser Ala Arg Ser Val Cys Thr Leu Phe Lys Ser Lys Ala
 165 170 175
 His

40 (2) INFORMATION FOR SEQ ID NO:63:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1041 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

5 (ix) FEATURE:

(A) NAME/KEY: Coding Sequence
 (B) LOCATION: 162...896
 (D) OTHER INFORMATION:

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

	AAAAGGATAC AATAAATACA AAAAATGAAA TTAAATAAA TAGGAATTTA ATGAGAATTT	60
	TTTTTGTTAT TATGGGACTT GTGTTTTTTG GTTGCACCAG TAAGGTGCAT GAGATGAAAA	120
	AAAGCCCTTG CACATTGTTA TGAAAACAGG TTAAATCTCG C ATG AAA GAA AAG CCT	176
	Met Lys Glu Lys Pro	
15	1 5	
	TTC AAT AGC GAG CAG TTG ATC TAT TTA GAA GAG CTT TTA AAC CAC CAA	224
	Phe Asn Ser Glu Gln Leu Ile Tyr Leu Glu Glu Leu Leu Asn His Gln	
	10 15 20	
	GAA AAG CAT TTA GAA AAC AAG CTT TCT GGT TTT TCG GTG AAT GAT TTG	272
20	Glu Lys His Leu Glu Asn Lys Leu Ser Gly Phe Ser Val Asn Asp Leu	
	25 30 35	
	GAC ATG CAA AGC GTG TTC AGA CTG GAG AGG AAC CGC TTG AAA ATC GCT	320
	Asp Met Gln Ser Val Phe Arg Leu Glu Arg Asn Arg Leu Lys Ile Ala	
	40 45 50	
25	TAT AAA CTC TTA GGC TTG ATG AGT TTT ATC GCT CTT GTT TTA GCG ATC	368
	Tyr Lys Leu Leu Gly Leu Met Ser Phe Ile Ala Leu Val Leu Ala Ile	
	55 60 65	
	GTG TTA ATC AGT GTT CTG CCC TTA CAA AAA ACC GAA CAC CAT TTC GTG	416
	Val Leu Ile Ser Val Leu Pro Leu Gln Lys Thr Glu His His Phe Val	
30	70 75 80 85	
	GAT TTT TTA AAT CAG GAC AAG CAT TAC GCC ATT ATC CAA AGA GCG GAT	464
	Asp Phe Leu Asn Gln Asp Lys His Tyr Ala Ile Ile Gln Arg Ala Asp	
	90 95 100	
	AAA AGC ATT TCC AGT AAT GAA GCG TTG GCT CGT TCG CTC ATT GGG GCG	512
35	Lys Ser Ile Ser Ser Asn Glu Ala Leu Ala Arg Ser Leu Ile Gly Ala	
	105 110 115	
	TAT GTG TTA AAC CGA GAG AGT ATT AAC CGC ATT GAC GAT AAA TCG CGC	560
	Tyr Val Leu Asn Arg Glu Ser Ile Asn Arg Ile Asp Asp Lys Ser Arg	

120

125

130

TAT GAA TTG GTG CGC TTG CAA AGC AGT TCT AAA GTG TGG CAA CGC TTT 608
 Tyr Glu Leu Val Arg Leu Gln Ser Ser Ser Lys Val Trp Gln Arg Phe
 135 140 145

5 GAA GAT TTG ATT AAA ACC CAA AAC AGC ATT TAT GTG CAA AGC CAT TTG 656
 Glu Asp Leu Ile Lys Thr Gln Asn Ser Ile Tyr Val Gln Ser His Leu
 150 155 160 165

GAA AGA GAA GTC CAT ATC GTC AAT ATT GCG ATC TAT CAG CAA GAC AAT 704
 Glu Arg Glu Val His Ile Val Asn Ile Ala Ile Tyr Gln Gln Asp Asn
 10 170 175 180

AAC CCC ATT GCG AGC GTC TCC ATT GCG GCT AAA CTT TTG AAC GAA AAC 752
 Asn Pro Ile Ala Ser Val Ser Ile Ala Ala Lys Leu Leu Asn Glu Asn
 185 190 195

AAG TTG GTG TAT GAA AAG CGT TAT AAA ATC GTA TTG AGT TAT TTG TTT 800
 15 Lys Leu Val Tyr Glu Lys Arg Tyr Lys Ile Val Leu Ser Tyr Leu Phe
 200 205 210

GAC ACC CCG GAT TTT GAT TAC GCT TCC ATG CCT AAA AAC CCT ACC GGA 848
 Asp Thr Pro Asp Phe Asp Tyr Ala Ser Met Pro Lys Asn Pro Thr Gly
 215 220 225

20 TTT AAA ATC ACC CGT TAC AGC ATC ACT GAA ATC ACT AAT AGG GGT GAT T 897
 Phe Lys Ile Thr Arg Tyr Ser Ile Thr Glu Ile Thr Asn Arg Gly Asp
 230 235 240 245

GATGCGTAAG GTTTTATACG CTCTTGTTGGG CTTTTTGTGTTG GCTTTTAGCG CTTTAAAGC 957
 CGATGATTTT TTAGAAGAAG CGAACGAAAC AGCCCCGGCG CATTTAAACC ACCCTATGCA 1017
 25 GGATTTAAAC GCCATTCAAG GGAG 1041

(2) INFORMATION FOR SEQ ID NO:64:

(i) SEQUENCE CHARACTERISTICS:

- 30 (A) LENGTH: 245 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

35 Met Lys Glu Lys Pro Phe Asn Ser Glu Gln Leu Ile Tyr Leu Glu Glu
 1 5 10 15
 Leu Leu Asn His Gln Glu Lys His Leu Glu Asn Lys Leu Ser Gly Phe

		20		25		30	
	Ser	Val	Asn	Asp	Leu	Asp	Met
		35		40		45	
5	Arg	Leu	Lys	Ile	Ala	Tyr	Lys
		50		55		60	
	Leu	Val	Leu	Ala	Ile	Val	Leu
		65		70		75	
	Glu	His	His	Phe	Val	Asp	Phe
			85		90		95
10	Ile	Gln	Arg	Ala	Asp	Lys	Ser
			100		105		110
	Ser	Leu	Ile	Gly	Ala	Tyr	Val
		115		120		125	
	Asp	Asp	Lys	Ser	Arg	Tyr	Glu
15		130		135		140	
	Val	Trp	Gln	Arg	Phe	Glu	Asp
		145		150		155	
	Val	Gln	Ser	His	Leu	Glu	Arg
			165		170		175
20	Tyr	Gln	Gln	Asp	Asn	Asn	Pro
			180		185		190
	Leu	Leu	Asn	Glu	Asn	Lys	Leu
		195		200		205	
	Leu	Ser	Tyr	Leu	Phe	Asp	Thr
25		210		215		220	
	Lys	Asn	Pro	Thr	Gly	Phe	Lys
		225		230		235	
	Thr	Asn	Arg	Gly	Asp		
			245				

30 (2) INFORMATION FOR SEQ ID NO:65:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2059 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

35

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 183...1961
- (D) OTHER INFORMATION:

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

GATATTTGTT	TTGTTGGGGG	TTAGGTTTTT	GTTTAAGAAA	GTTTTTTAAA	ACTAAAGAAG	60
CGCTTAAAC	AGAACCTTTT	GTTTTTTAGG	TTTTATTTTT	TACTTTGGCT	TGTTTTCAAA	120
AGTCATTTT	ATTCTAAAA	ATAGTCTATA	ATGCTCGCAA	GAGATATTTT	TTAAGGTTAT	180

CA ATG AAA GCT ATA AAA ATA CTT TTT ATA ATG ACA CTC AGT TTA AAC 227
 Met Lys Ala Ile Lys Ile Leu Phe Ile Met Thr Leu Ser Leu Asn
 1 5 10 15

GCT ATC AGC GTG AAT AGG GCG TTG TTT GAT TTA AAA GAT TCG CAA TTA 275
 5 Ala Ile Ser Val Asn Arg Ala Leu Phe Asp Leu Lys Asp Ser Gln Leu
 20 25 30

AAA GGG GAA TTA ACG CCA AAA ATA GTG AAT TTT GGG GGT TAT AAA AGC 323
 Lys Gly Glu Leu Thr Pro Lys Ile Val Asn Phe Gly Gly Tyr Lys Ser
 35 40 45

10 AGC ACT GAA GAG TGG GGG GCT ACG GCT TTA AAC TAT ATC AAT GCG GCT 371
 Ser Thr Glu Glu Trp Gly Ala Thr Ala Leu Asn Tyr Ile Asn Ala Ala
 50 55 60

AAT GGC GAT GCG AAA AAA TTC AGC ACT CTA GTG GAA AAA ATG CGT TTT 419
 Asn Gly Asp Ala Lys Lys Phe Ser Thr Leu Val Glu Lys Met Arg Phe
 15 65 70 75

AAC TCC GGT ATA TTG GGG AAT TTA AGA GTG CAT GCA CGT TTG AGG CAA 467
 Asn Ser Gly Ile Leu Gly Asn Leu Arg Val His Ala Arg Leu Arg Gln
 80 85 90 95

GCC CTA AAA TTG CAA AAG AAT TTG AAA TAT TGC CTT AAA ATC ATC GCT 515
 20 Ala Leu Lys Leu Gln Lys Asn Leu Lys Tyr Cys Leu Lys Ile Ile Ala
 100 105 110

AGG GAT TCT TTT TAT AGC TAC CGC ACC GGT ATT TAT ATC CCC TTA GGC 563
 Arg Asp Ser Phe Tyr Ser Tyr Arg Thr Gly Ile Tyr Ile Pro Leu Gly
 115 120 125

25 ATT TCT TTA AAA GAT CAA AAA ACG GCT CAA AAA ATG CTC GCT GAT TTG 611
 Ile Ser Leu Lys Asp Gln Lys Thr Ala Gln Lys Met Leu Ala Asp Leu
 130 135 140

AGC GTG GTA GGG GCG TAT CTT AAA AAA CAA CAA GAG AAT GAA AAG GCT 659
 30 Ser Val Val Gly Ala Tyr Leu Lys Lys Gln Gln Glu Asn Glu Lys Ala
 145 150 155

CAA AGC CCT TAT TAC AGA AAC AAC AAC TAT TAC AAC TCT TAC TAT AGC 707
 Gln Ser Pro Tyr Tyr Arg Asn Asn Asn Tyr Tyr Asn Ser Tyr Tyr Ser
 160 165 170 175

CCT TAT TAC GGA ATG TAT GGT ATG TAT GGC ATG GGC ATG TAT GGA ATG 755
 35 Pro Tyr Tyr Gly Met Tyr Gly Met Tyr Gly Met Gly Met Tyr Gly Met
 180 185 190

TAT GGC ATG GGC ATG TAT GAT TTT TAT GAC TTT TAT GAT GGC ATG TAT 803
 Tyr Gly Met Gly Met Tyr Asp Phe Tyr Asp Phe Tyr Asp Gly Met Tyr

	195	200	205	
	GGA TTC TAC CCT AAC ATG TTT	TTC ATG ATG CAA GTT CAA GAT TAC TTG		851
	Gly Phe Tyr Pro Asn Met Phe	Phe Met Met Gln Val Gln Asp Tyr Leu		
	210	215	220	
5	ATG TTA GAA AAT TAC ATG TAT	GCG CTC GAT CAA GAA GAG ATT TTA GAT		899
	Met Leu Glu Asn Tyr Met Tyr	Ala Leu Asp Gln Glu Glu Ile Leu Asp		
	225	230	235	
	CAT GAC GCT TCT ACT GAC CAA CTT	GAT ACG CCT ACT GAT GAT GAC AAA		947
10	His Asp Ala Ser Thr Asp Gln	Leu Asp Thr Pro Thr Asp Asp Asp Lys		
	240	245	250	255
	GAC GAT AAA GAC GAT AAA TCC TTA	CAG CAG GCA AAT CTT ATG AAC TTT		995
	Asp Asp Lys Asp Asp Lys Ser	Leu Gln Gln Ala Asn Leu Met Asn Phe		
	260	265	270	
	TAT CGT GAT CCC AAA TTC AGC AAA	GGC ATT CAA ACC AAC CGC TTG AAT		1043
15	Tyr Arg Asp Pro Lys Phe Ser	Lys Gly Ile Gln Thr Asn Arg Leu Asn		
	275	280	285	
	AGC GCT TTA GTC AAT TTA GAC AAC	AGT CGC ATG CTC AAA GAC AAT TCG		1091
	Ser Ala Leu Val Asn Leu Asp	Asn Ser Arg Met Leu Lys Asp Asn Ser		
	290	295	300	
20	CTT TTC CAC ACT AAA GCC ATG CCC	ACT AAA AGC GTG GAT GCG ATA ACT		1139
	Leu Phe His Thr Lys Ala Met	Pro Thr Lys Ser Val Asp Ala Ile Thr		
	305	310	315	
	TCT CAA GCC AAA GAG CTT AAC CAT	TTA GTG GGG CAA ATC AAA GAA ATG		1187
	Ser Gln Ala Lys Glu Leu Asn	His Leu Val Gly Gln Ile Lys Glu Met		
25	320	325	330	335
	AAG CAA GAC GGG GCG AGT CCT AGT	AAG ATT GAT TCA GTT GTC AAT AAA		1235
	Lys Gln Asp Gly Ala Ser Pro	Ser Lys Ile Asp Ser Val Val Asn Lys		
	340	345	350	
	GCT ATG GAA GTG AGG GAC AAG CTA	GAC AAT AAT CTC AAC CAA CTA GAC		1283
30	Ala Met Glu Val Arg Asp Lys	Leu Asp Asn Asn Leu Asn Gln Leu Asp		
	355	360	365	
	AAT GAC TTA AAA GAT CAA AAA GGG	CTT TCA AGC GAG CAA CAA GCT CAA		1331
	Asn Asp Leu Lys Asp Gln Lys	Gly Leu Ser Ser Glu Gln Gln Ala Gln		
	370	375	380	
35	GTG GAT AAA GCC CTA GAC AGC GTG	CAA CAA TTA AGC CAT AGC AGC GAT		1379
	Val Asp Lys Ala Leu Asp Ser	Val Gln Gln Leu Ser His Ser Ser Asp		
	385	390	395	

	GTG GTG GGG AAT TAT TTA GAC GGG AGT TTG AAA ATT GAT GGC GAT GAT	1427
	Val Val Gly Asn Tyr Leu Asp Gly Ser Leu Lys Ile Asp Gly Asp Asp	
	400 405 410 415	
5	AGA GAT GAT TTG AAT GAT GCG ATG AAT AAC CCT ATG CAA CAA CCC GTG	1475
	Arg Asp Asp Leu Asn Asp Ala Met Asn Asn Pro Met Gln Gln Pro Val	
	420 425 430	
	CAA CAA ACG CCT ACT AGC AAC ATG GCC GAC ACC CAT GCA AAT GAC AGC	1523
	Gln Gln Thr Pro Thr Ser Asn Met Ala Asp Thr His Ala Asn Asp Ser	
	435 440 445	
10	AAG GAT CAA GGG AGT AAC GCG CTC ATA AAC CCT AAC AGC GCC ACT AAC	1571
	Lys Asp Gln Gly Ser Asn Ala Leu Ile Asn Pro Asn Ser Ala Thr Asn	
	450 455 460	
	GCC GAC GAC ACT CAC ACT GAC GAT ACT CAC ACT GAC ACT AAC ACC ACA	1619
	Ala Asp Asp Thr His Thr Asp Asp Thr His Thr Asp Thr Asn Thr Thr	
15	465 470 475	
	AAC GAT GCT AGC ACC ACT GAC ACC CCC ACT GAC GAT AAA GAT GCT AGC	1667
	Asn Asp Ala Ser Thr Thr Asp Thr Pro Thr Asp Asp Lys Asp Ala Ser	
	480 485 490 495	
20	GGC TTG AAC AAT ACC GGC GAT ATG AAT AAC ACG GAT ACC GGC AAC ACG	1715
	Gly Leu Asn Asn Thr Gly Asp Met Asn Asn Thr Asp Thr Gly Asn Thr	
	500 505 510	
	GAC ACC GGC AAT ACG GAT ACC GGT AAC ACT GAT GAT ATG AGC AAC ATG	1763
	Asp Thr Gly Asn Thr Asp Thr Gly Asn Thr Asp Asp Met Ser Asn Met	
	515 520 525	
25	AAC AAC GGC AAC GAT GAT ACG GGT AAC GCT AAT GAC GAC ATG AGC AAC	1811
	Asn Asn Gly Asn Asp Asp Thr Gly Asn Ala Asn Asp Asp Met Ser Asn	
	530 535 540	
	GGC AAC GAC ATG GGC GAT GAT TTG AAC AAC GCG AAC GAT ATG AAC GAC	1859
	Gly Asn Asp Met Gly Asp Asp Leu Asn Asn Ala Asn Asp Met Asn Asp	
30	545 550 555	
	GAC ATG GGT AAT GGC AAC GAT GAC ATG GGC GAT ATG GGG GAT ATG AAC	1907
	Asp Met Gly Asn Gly Asn Asp Asp Met Gly Asp Met Gly Asp Met Asn	
	560 565 570 575	
35	GAC GAT ATG GGT GGC GAT ATG GGA GAC ATG GGG GAT ATG GGC GAT ATG	1955
	Asp Asp Met Gly Gly Asp Met Gly Asp Met Gly Asp Met Gly Asp Met	
	580 585 590	
	GGG AAT TGAGATTAAAC CCCAATATCA AAGAGTGATA GCCAAACTT TAAGGAATAT TT	2013
	Gly Asn	

TTATAGTAAA AACGATTCTT TTAAGGTAAT AGGGGGGATA TTTTGC

2059

(2) INFORMATION FOR SEQ ID NO:66:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 593 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

	Met	Lys	Ala	Ile	Lys	Ile	Leu	Phe	Ile	Met	Thr	Leu	Ser	Leu	Asn	Ala	
	1				5					10					15		
	Ile	Ser	Val	Asn	Arg	Ala	Leu	Phe	Asp	Leu	Lys	Asp	Ser	Gln	Leu	Lys	
				20					25					30			
15	Gly	Glu	Leu	Thr	Pro	Lys	Ile	Val	Asn	Phe	Gly	Gly	Tyr	Lys	Ser	Ser	
				35				40					45				
	Thr	Glu	Glu	Trp	Gly	Ala	Thr	Ala	Leu	Asn	Tyr	Ile	Asn	Ala	Ala	Asn	
				50				55				60					
	Gly	Asp	Ala	Lys	Lys	Phe	Ser	Thr	Leu	Val	Glu	Lys	Met	Arg	Phe	Asn	
20	65					70					75				80		
	Ser	Gly	Ile	Leu	Gly	Asn	Leu	Arg	Val	His	Ala	Arg	Leu	Arg	Gln	Ala	
					85					90					95		
	Leu	Lys	Leu	Gln	Lys	Asn	Leu	Lys	Tyr	Cys	Leu	Lys	Ile	Ile	Ala	Arg	
				100					105					110			
25	Asp	Ser	Phe	Tyr	Ser	Tyr	Arg	Thr	Gly	Ile	Tyr	Ile	Pro	Leu	Gly	Ile	
				115				120					125				
	Ser	Leu	Lys	Asp	Gln	Lys	Thr	Ala	Gln	Lys	Met	Leu	Ala	Asp	Leu	Ser	
				130				135				140					
	Val	Val	Gly	Ala	Tyr	Leu	Lys	Lys	Gln	Gln	Glu	Asn	Glu	Lys	Ala	Gln	
30	145					150					155				160		
	Ser	Pro	Tyr	Tyr	Arg	Asn	Asn	Asn	Tyr	Tyr	Asn	Ser	Tyr	Tyr	Ser	Pro	
					165					170					175		
	Tyr	Tyr	Gly	Met	Tyr	Gly	Met	Tyr	Gly	Met	Gly	Met	Tyr	Gly	Met	Tyr	
				180					185					190			
35	Gly	Met	Gly	Met	Tyr	Asp	Phe	Tyr	Asp	Phe	Tyr	Asp	Gly	Met	Tyr	Gly	
				195				200					205				
	Phe	Tyr	Pro	Asn	Met	Phe	Phe	Met	Met	Gln	Val	Gln	Asp	Tyr	Leu	Met	
				210				215					220				
	Leu	Glu	Asn	Tyr	Met	Tyr	Ala	Leu	Asp	Gln	Glu	Glu	Ile	Leu	Asp	His	
40	225					230					235				240		
	Asp	Ala	Ser	Thr	Asp	Gln	Leu	Asp	Thr	Pro	Thr	Asp	Asp	Asp	Lys	Asp	
					245					250					255		
	Asp	Lys	Asp	Asp	Lys	Ser	Leu	Gln	Gln	Ala	Asn	Leu	Met	Asn	Phe	Tyr	
				260					265					270			

	Arg	Asp	Pro	Lys	Phe	Ser	Lys	Gly	Ile	Gln	Thr	Asn	Arg	Leu	Asn	Ser
			275					280					285			
	Ala	Leu	Val	Asn	Leu	Asp	Asn	Ser	Arg	Met	Leu	Lys	Asp	Asn	Ser	Leu
		290					295					300				
5	Phe	His	Thr	Lys	Ala	Met	Pro	Thr	Lys	Ser	Val	Asp	Ala	Ile	Thr	Ser
	305					310						315				320
	Gln	Ala	Lys	Glu	Leu	Asn	His	Leu	Val	Gly	Gln	Ile	Lys	Glu	Met	Lys
					325					330					335	
	Gln	Asp	Gly	Ala	Ser	Pro	Ser	Lys	Ile	Asp	Ser	Val	Val	Asn	Lys	Ala
10				340					345					350		
	Met	Glu	Val	Arg	Asp	Lys	Leu	Asp	Asn	Asn	Leu	Asn	Gln	Leu	Asp	Asn
			355					360					365			
	Asp	Leu	Lys	Asp	Gln	Lys	Gly	Leu	Ser	Ser	Glu	Gln	Gln	Ala	Gln	Val
		370					375					380				
15	Asp	Lys	Ala	Leu	Asp	Ser	Val	Gln	Gln	Leu	Ser	His	Ser	Ser	Asp	Val
	385					390					395					400
	Val	Gly	Asn	Tyr	Leu	Asp	Gly	Ser	Leu	Lys	Ile	Asp	Gly	Asp	Asp	Arg
				405					410					415		
	Asp	Asp	Leu	Asn	Asp	Ala	Met	Asn	Asn	Pro	Met	Gln	Gln	Pro	Val	Gln
20				420					425					430		
	Gln	Thr	Pro	Thr	Ser	Asn	Met	Ala	Asp	Thr	His	Ala	Asn	Asp	Ser	Lys
		435						440					445			
	Asp	Gln	Gly	Ser	Asn	Ala	Leu	Ile	Asn	Pro	Asn	Ser	Ala	Thr	Asn	Ala
		450					455					460				
25	Asp	Asp	Thr	His	Thr	Asp	Asp	Thr	His	Thr	Asp	Thr	Asn	Thr	Thr	Asn
	465					470					475					480
	Asp	Ala	Ser	Thr	Thr	Asp	Thr	Pro	Thr	Asp	Asp	Lys	Asp	Ala	Ser	Gly
				485					490					495		
	Leu	Asn	Asn	Thr	Gly	Asp	Met	Asn	Asn	Thr	Asp	Thr	Gly	Asn	Thr	Asp
30				500					505					510		
	Thr	Gly	Asn	Thr	Asp	Thr	Gly	Asn	Thr	Asp	Asp	Met	Ser	Asn	Met	Asn
		515						520					525			
	Asn	Gly	Asn	Asp	Asp	Thr	Gly	Asn	Ala	Asn	Asp	Asp	Met	Ser	Asn	Gly
		530					535					540				
35	Asn	Asp	Met	Gly	Asp	Asp	Leu	Asn	Asn	Ala	Asn	Asp	Met	Asn	Asp	Asp
	545					550					555					560
	Met	Gly	Asn	Gly	Asn	Asp	Asp	Met	Gly	Asp	Met	Gly	Asp	Met	Asn	Asp
				565					570					575		
	Asp	Met	Gly	Gly	Asp	Met	Gly	Asp	Met	Gly	Asp	Met	Gly	Asp	Met	Gly
40				580				585					590			
	Asn															

(2) INFORMATION FOR SEQ ID NO:67:

(i) SEQUENCE CHARACTERISTICS:

- 45 (A) LENGTH: 1527 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence

(B) LOCATION: 112...1461

(D) OTHER INFORMATION:

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

AATGAGCGAT TTGAAAGATT TTGTCAATAA AACTTCAAGC CCTTTAAATG CGAATTGATT 60
TTCTTATATT ATGATTACGA TTTATCAATT TAAAACATTT GGAGAAAGAC A ATG AGT 117
Met Ser
1

10

ATG GAA TTT GAT GCT GTT ATT ATT GGA GGT GGG GTT TCA GGG TGC GCG 165
Met Glu Phe Asp Ala Val Ile Ile Gly Gly Gly Val Ser Gly Cys Ala
5 10 15

15

ACC TTT TAT ACT TTG AGC GAA TAC AGC TCT TTA AAG CGC GTG GCT ATC 213
Thr Phe Tyr Thr Leu Ser Glu Tyr Ser Ser Leu Lys Arg Val Ala Ile
20 25 30

GTG GAA AAA TGC TCT AAA TTG GCT CAA ATC AGC TCC AGC GCT AAA GCT 261
Val Glu Lys Cys Ser Lys Leu Ala Gln Ile Ser Ser Ser Ala Lys Ala
35 40 45 50

20

AAT TCG CAA ACC ATT CAT GAT GGC TCT ATT GAA ACG AAT TAC ACT CCC 309
Asn Ser Gln Thr Ile His Asp Gly Ser Ile Glu Thr Asn Tyr Thr Pro
55 60 65

25

GAA AAA GCT AAA AAA GTG CGT TTG AGC GCT TAT AAG ACC AGG CAA TAC 357
Glu Lys Ala Lys Lys Val Arg Leu Ser Ala Tyr Lys Thr Arg Gln Tyr
70 75 80

GCT TTG AAT AAA GGC TTG CAA AAT GAA GTG ATT TTT GAA ACC CAG AAA 405
Ala Leu Asn Lys Gly Leu Gln Asn Glu Val Ile Phe Glu Thr Gln Lys
85 90 95

30

ATG GCT ATA GGC GTG GGC GAT GAA GAA TGC GAG TTC ATG AAA AAA CGC 453
Met Ala Ile Gly Val Gly Asp Glu Glu Cys Glu Phe Met Lys Lys Arg
100 105 110

TAC GAA TCT TTT AAA GAA ATC TTT GTG GGG TTA GAA GAA TTT GAC AAG 501
Tyr Glu Ser Phe Lys Glu Ile Phe Val Gly Leu Glu Glu Phe Asp Lys
115 120 125 130

35

CAA AAG ATT AAA GAA TTA GAG CCT AAT GTG ATT TTA GGG GCT AAT GGC 549
Gln Lys Ile Lys Glu Leu Glu Pro Asn Val Ile Leu Gly Ala Asn Gly
135 140 145

	ATA GAC AGG CAT GAA AAC ATT ATC GGG CAT GGG TAT AGA AAG GAT TGG	597
	Ile Asp Arg His Glu Asn Ile Ile Gly His Gly Tyr Arg Lys Asp Trp	
	150 155 160	
5	AGC ACC ATG AAT TTT GCG AAG TTG AGT GAA AAC TTC GTT GAA GAA GCC	645
	Ser Thr Met Asn Phe Ala Lys Leu Ser Glu Asn Phe Val Glu Glu Ala	
	165 170 175	
	CTA AAA TTA AAG CCT AAC AAC CAG GTG TTT TTG AAT TTC AAA GTG AAA	693
	Leu Lys Leu Lys Pro Asn Asn Gln Val Phe Leu Asn Phe Lys Val Lys	
	180 185 190	
10	AAG ATT GAA AAA CGC AAC GAC ACT TAC GCC GTA ATT TCA GAA GAC GCT	741
	Lys Ile Glu Lys Arg Asn Asp Thr Tyr Ala Val Ile Ser Glu Asp Ala	
	195 200 205 210	
	GAA GAA GTG TAT GCT AAA TTC GTG CTG GTC AAT GCC GGC TCT TAC GCT	789
	Glu Glu Val Tyr Ala Lys Phe Val Leu Val Asn Ala Gly Ser Tyr Ala	
15	215 220 225	
	TTG CCT TTG GCT CAG AGC ATG GGC TAT GGC CTA GAT TTA GGG TGC TTG	837
	Leu Pro Leu Ala Gln Ser Met Gly Tyr Gly Leu Asp Leu Gly Cys Leu	
	230 235 240	
20	CCT GTG GCG GGC AGC TTT TAT TTT GTG CCG GAT TTA TTA AGG GGT AAG	885
	Pro Val Ala Gly Ser Phe Tyr Phe Val Pro Asp Leu Leu Arg Gly Lys	
	245 250 255	
	GTT TAT ACC GTT CAA AAC CCC AAA CTC CCT TTT GCA GCC GTG CAT GGC	933
	Val Tyr Thr Val Gln Asn Pro Lys Leu Pro Phe Ala Ala Val His Gly	
	260 265 270	
25	GAC CCT GAT GCC GTC ATT AAA GGA AAA ACA CGA ATC GGG CCT ACC GCT	981
	Asp Pro Asp Ala Val Ile Lys Gly Lys Thr Arg Ile Gly Pro Thr Ala	
	275 280 285 290	
	TTA ACG ATG CCT AAA TTA GAA CGC AAC AAA TGT TGG CTT AAG GGC ATT	1029
	Leu Thr Met Pro Lys Leu Glu Arg Asn Lys Cys Trp Leu Lys Gly Ile	
30	295 300 305	
	AGC TTG GAA TTG TTG AAA ATG GAT TTG AAT AAA GAT GTG TTT AAA ATT	1077
	Ser Leu Glu Leu Leu Lys Met Asp Leu Asn Lys Asp Val Phe Lys Ile	
	310 315 320	
	GCG TTT GAT TTG ATG AGC GAT AAA GAA ATC CGA AAT TAT GTG TTT AAA	1125
35	Ala Phe Asp Leu Met Ser Asp Lys Glu Ile Arg Asn Tyr Val Phe Lys	
	325 330 335	
	AAC ATG GTT TTT GAA TTG CCC ATT ATC GGT AAA AGG AAA TTT TTA AAA	1173
	Asn Met Val Phe Glu Leu Pro Ile Ile Gly Lys Arg Lys Phe Leu Lys	

340

345

350

GAC GCT CAA AAA ATC ATC CCC TCT CTT AGC CTA GAA GAT CTA GAA TAC 1221
 Asp Ala Gln Lys Ile Ile Pro Ser Leu Ser Leu Glu Asp Leu Glu Tyr
 355 360 365 370

5 GCT CAT GGT TTT GGT GAA GTG CGC CCG CAA GTT TTA GAC AGA ACC AAG 1269
 Ala His Gly Phe Gly Glu Val Arg Pro Gln Val Leu Asp Arg Thr Lys
 375 380 385

CGA AAA CTG GAA TTA GGC GAA AAA AAG ATT TGC ACC CAT AAA GGC ATC 1317
 Arg Lys Leu Glu Leu Gly Glu Lys Lys Ile Cys Thr His Lys Gly Ile
 10 390 395 400

ACT TTT AAC ATG ACC CCT TCT CCA GGC GCG ACG AGT TGT TTG CAA AAC 1365
 Thr Phe Asn Met Thr Pro Ser Pro Gly Ala Thr Ser Cys Leu Gln Asn
 405 410 415

GCC CTT GTG GAT TCC CAA GAA ATC GCT GCG TAT TTG GGC GAG AGC TTT 1413
 15 Ala Leu Val Asp Ser Gln Glu Ile Ala Ala Tyr Leu Gly Glu Ser Phe
 420 425 430

GAA TTA GAA CGC TTT TAT AAA GAT TTA TCC CCA GAA GAA TTG GAA AAT T 1462
 Glu Leu Glu Arg Phe Tyr Lys Asp Leu Ser Pro Glu Glu Leu Glu Asn
 435 440 445 450

20 AAAAAACGCAT GCAAAAAGAA CAAGAAGCCC AAGAAATCGC TAAAAAGCC GTTAAAATCG 1522
 TGTTT 1527

(2) INFORMATION FOR SEQ ID NO:68:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 450 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

Met Ser Met Glu Phe Asp Ala Val Ile Ile Gly Gly Gly Val Ser Gly
 1 5 10 15
 Cys Ala Thr Phe Tyr Thr Leu Ser Glu Tyr Ser Ser Leu Lys Arg Val
 20 25 30
 35 Ala Ile Val Glu Lys Cys Ser Lys Leu Ala Gln Ile Ser Ser Ser Ala
 35 40 45
 Lys Ala Asn Ser Gln Thr Ile His Asp Gly Ser Ile Glu Thr Asn Tyr
 50 55 60

	Thr	Pro	Glu	Lys	Ala	Lys	Lys	Val	Arg	Leu	Ser	Ala	Tyr	Lys	Thr	Arg
	65					70					75					80
	Gln	Tyr	Ala	Leu	Asn	Lys	Gly	Leu	Gln	Asn	Glu	Val	Ile	Phe	Glu	Thr
				85					90						95	
5	Gln	Lys	Met	Ala	Ile	Gly	Val	Gly	Asp	Glu	Glu	Cys	Glu	Phe	Met	Lys
				100					105					110		
	Lys	Arg	Tyr	Glu	Ser	Phe	Lys	Glu	Ile	Phe	Val	Gly	Leu	Glu	Glu	Phe
				115				120					125			
	Asp	Lys	Gln	Lys	Ile	Lys	Glu	Leu	Glu	Pro	Asn	Val	Ile	Leu	Gly	Ala
10		130					135					140				
	Asn	Gly	Ile	Asp	Arg	His	Glu	Asn	Ile	Ile	Gly	His	Gly	Tyr	Arg	Lys
	145					150					155					160
	Asp	Trp	Ser	Thr	Met	Asn	Phe	Ala	Lys	Leu	Ser	Glu	Asn	Phe	Val	Glu
					165					170					175	
15	Glu	Ala	Leu	Lys	Leu	Lys	Pro	Asn	Asn	Gln	Val	Phe	Leu	Asn	Phe	Lys
				180					185					190		
	Val	Lys	Lys	Ile	Glu	Lys	Arg	Asn	Asp	Thr	Tyr	Ala	Val	Ile	Ser	Glu
				195				200					205			
	Asp	Ala	Glu	Glu	Val	Tyr	Ala	Lys	Phe	Val	Leu	Val	Asn	Ala	Gly	Ser
20		210					215					220				
	Tyr	Ala	Leu	Pro	Leu	Ala	Gln	Ser	Met	Gly	Tyr	Gly	Leu	Asp	Leu	Gly
	225					230					235					240
	Cys	Leu	Pro	Val	Ala	Gly	Ser	Phe	Tyr	Phe	Val	Pro	Asp	Leu	Leu	Arg
					245					250					255	
25	Gly	Lys	Val	Tyr	Thr	Val	Gln	Asn	Pro	Lys	Leu	Pro	Phe	Ala	Ala	Val
				260					265					270		
	His	Gly	Asp	Pro	Asp	Ala	Val	Ile	Lys	Gly	Lys	Thr	Arg	Ile	Gly	Pro
			275					280					285			
	Thr	Ala	Leu	Thr	Met	Pro	Lys	Leu	Glu	Arg	Asn	Lys	Cys	Trp	Leu	Lys
30		290					295					300				
	Gly	Ile	Ser	Leu	Glu	Leu	Leu	Lys	Met	Asp	Leu	Asn	Lys	Asp	Val	Phe
	305					310					315					320
	Lys	Ile	Ala	Phe	Asp	Leu	Met	Ser	Asp	Lys	Glu	Ile	Arg	Asn	Tyr	Val
					325					330					335	
35	Phe	Lys	Asn	Met	Val	Phe	Glu	Leu	Pro	Ile	Ile	Gly	Lys	Arg	Lys	Phe
				340					345					350		
	Leu	Lys	Asp	Ala	Gln	Lys	Ile	Ile	Pro	Ser	Leu	Ser	Leu	Glu	Asp	Leu
			355					360					365			
	Glu	Tyr	Ala	His	Gly	Phe	Gly	Glu	Val	Arg	Pro	Gln	Val	Leu	Asp	Arg
40		370					375					380				
	Thr	Lys	Arg	Lys	Leu	Glu	Leu	Gly	Glu	Lys	Lys	Ile	Cys	Thr	His	Lys
	385					390					395					400
	Gly	Ile	Thr	Phe	Asn	Met	Thr	Pro	Ser	Pro	Gly	Ala	Thr	Ser	Cys	Leu
					405					410					415	
45	Gln	Asn	Ala	Leu	Val	Asp	Ser	Gln	Glu	Ile	Ala	Ala	Tyr	Leu	Gly	Glu
				420					425					430		
	Ser	Phe	Glu	Leu	Glu	Arg	Phe	Tyr	Lys	Asp	Leu	Ser	Pro	Glu	Glu	Leu
			435					440					445			
	Glu	Asn														
50		450														

(2) INFORMATION FOR SEQ ID NO:69:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 653 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
(B) LOCATION: 63...590
(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

CTAGATTTAA TTTTAAAGTT ATATAATTAA ACCACAAAAT CCTTTTTTAA AAGAAACTAA 60
GC ATG CCA AAA CCC AAG AAA AAC ACC CTC CCC TGT AGC CTT TCT GTC 107
15 Met Pro Lys Pro Lys Lys Asn Thr Leu Pro Cys Ser Leu Ser Val
1 5 10 15
AAA ATG TCT TAT TTC ATG CGC TTT CTC ATT AAA TGG CGC ACC CGC TCT 155
Lys Met Ser Tyr Phe Met Arg Phe Leu Ile Lys Trp Arg Thr Arg Ser
20 25 30
20 TTA AGC CAT AAA ATG ATG ACT CTC ATT CAA ATC TTA AGC ATT CTG GCT 203
Leu Ser His Lys Met Met Thr Leu Ile Gln Ile Leu Ser Ile Leu Ala
35 40 45
TTA GCG AGC AAG GCC AGT GAA GAT TTA GAA GAG CAA CTC AAA AAA ATC 251
Leu Ala Ser Lys Ala Ser Glu Asp Leu Glu Glu Gln Leu Lys Lys Ile
25 50 55 60
AAA GAT TAC ATT TAT AGA ACC CTA AAC GCT AAA ATC GCA TCG GAT GTG 299
Lys Asp Tyr Ile Tyr Arg Thr Leu Asn Ala Lys Ile Ala Ser Asp Val
65 70 75
TAT AAC CGA GTG CTT ATT TTA GTG AAT GAA TAT TGC ACT AAT GAA GAA 347
Tyr Asn Arg Val Leu Ile Leu Val Asn Glu Tyr Cys Thr Asn Glu Glu
80 85 90 95
TTG TTT GAC AAA GAG AGC GTT AAA ATT TCA GAT TTA CTC ATT CAA GAC 395
Leu Phe Asp Lys Glu Ser Val Lys Ile Ser Asp Leu Leu Ile Gln Asp
100 105 110
35 ATT CAG CTT TAC GCT TTA GTG GAT GAA ATG CTT AAA GAA GAT AAA TAT 443
Ile Gln Leu Tyr Ala Leu Val Asp Glu Met Leu Lys Glu Asp Lys Tyr
115 120 125

CAA GTC CAG CAC ACC ATT TTA AAG GGC ATC ATC AAA CGC AAA TAC GAT 491
 Gln Val Gln His Thr Ile Leu Lys Gly Ile Ile Lys Arg Lys Tyr Asp
 130 135 140

GAA GCC TAC TCG CTC AAT AGC GAA GAC AGG ATT CTT TTA GAA TAC CAA 539
 5 Glu Ala Tyr Ser Leu Asn Ser Glu Asp Arg Ile Leu Leu Glu Tyr Gln
 145 150 155

GAA CGC TTG CTA GAA CAC TCA CAC GCG TCT TTT TCA AAT AAA AAA TTC 587
 Glu Arg Leu Leu Glu His Ser His Ala Ser Phe Ser Asn Lys Lys Phe
 160 165 170 175

10 AAA TGATTGAAA GCGTTACTTG CCCTGCTTTT TGGGCTTTTA TTGAAAAGG GCTTTA 646
 Lys

AAATGAG 653

(2) INFORMATION FOR SEQ ID NO:70:

(i) SEQUENCE CHARACTERISTICS:

- 15 (A) LENGTH: 176 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

- 20 (v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

Met Pro Lys Pro Lys Lys Asn Thr Leu Pro Cys Ser Leu Ser Val Lys
 1 5 10 15
 Met Ser Tyr Phe Met Arg Phe Leu Ile Lys Trp Arg Thr Arg Ser Leu
 20 25 30
 Ser His Lys Met Met Thr Leu Ile Gln Ile Leu Ser Ile Leu Ala Leu
 35 40 45
 Ala Ser Lys Ala Ser Glu Asp Leu Glu Glu Gln Leu Lys Lys Ile Lys
 50 55 60
 30 Asp Tyr Ile Tyr Arg Thr Leu Asn Ala Lys Ile Ala Ser Asp Val Tyr
 65 70 75 80
 Asn Arg Val Leu Ile Leu Val Asn Glu Tyr Cys Thr Asn Glu Glu Leu
 85 90 95
 Phe Asp Lys Glu Ser Val Lys Ile Ser Asp Leu Leu Ile Gln Asp Ile
 100 105 110
 35 Gln Leu Tyr Ala Leu Val Asp Glu Met Leu Lys Glu Asp Lys Tyr Gln
 115 120 125
 Val Gln His Thr Ile Leu Lys Gly Ile Ile Lys Arg Lys Tyr Asp Glu
 130 135 140
 40 Ala Tyr Ser Leu Asn Ser Glu Asp Arg Ile Leu Leu Glu Tyr Gln Glu

145	150	155	160
Arg Leu Leu Glu His Ser His Ala Ser Phe Ser Asn Lys Lys Phe Lys			
165	170	175	

(2) INFORMATION FOR SEQ ID NO:71:

5 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1840 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

10 (ix) FEATURE:
 (A) NAME/KEY: Coding Sequence
 (B) LOCATION: 91...1833
 (D) OTHER INFORMATION:

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

	AAGCGTTAAA TTCCAATCAA AAACCATCGT ATCGGTGTTA ATATTGTGTA AAAATTAATG	60
	TTATGAATCT CTTGTATTAA AAGGACTTCA ATG AAA AAA TTG GTT TTA GTC ATC	114
	Met Lys Lys Leu Val Leu Val Ile	
	1 5	
20	TTT TTA ACG CTA GCG CTT TCA ATA TCT GCA AAA GAA GTC AAA ATA GTG	162
	Phe Leu Thr Leu Ala Leu Ser Ile Ser Ala Lys Glu Val Lys Ile Val	
	10 15 20	
	TTT TTA GAA ACT TCA GAC ATT CAT GGG CGG CTT TTT TCG TAT GAT TAT	210
	Phe Leu Glu Thr Ser Asp Ile His Gly Arg Leu Phe Ser Tyr Asp Tyr	
25	25 30 35 40	
	GCG ATT GGC GAG CAA AAA CCC AAT AAC GGC TTG ACA AGG ATT GCG ACT	258
	Ala Ile Gly Glu Gln Lys Pro Asn Asn Gly Leu Thr Arg Ile Ala Thr	
	45 50 55	
30	TTA ATC AAA AAG CAA AGG GCT GAG AAT AAA AAT GTG GTT TTG ATT GAC	306
	Leu Ile Lys Lys Gln Arg Ala Glu Asn Lys Asn Val Val Leu Ile Asp	
	60 65 70	
	AGC GGG GAT TTG TTG CAA GGC AAT AGC GCG GAG TTG TTT AAT GAT GAG	354
	Ser Gly Asp Leu Leu Gln Gly Asn Ser Ala Glu Leu Phe Asn Asp Glu	
	75 80 85	
35	CCA ATT CAT CCG CTA GTT AGA GCT GAA AAC GAT TTG AAA TTT GAC ATT	402
	Pro Ile His Pro Leu Val Arg Ala Glu Asn Asp Leu Lys Phe Asp Ile	
	90 95 100	

	CGT GTG CTT GGC AAT CAC GAG TTT AAT TTC AGT AAA GAT TTT TTA GAA	450
	Arg Val Leu Gly Asn His Glu Phe Asn Phe Ser Lys Asp Phe Leu Glu	
	105 110 115 120	
5	AAG AAT ATT AAG GGG TTT AAT GGC GAT GTC ATG AAT GCG AAT ATC ATT	498
	Lys Asn Ile Lys Gly Phe Asn Gly Asp Val Met Asn Ala Asn Ile Ile	
	125 130 135	
	AAA ATT GCG GAC AAT AAG CCG TTT GTA AAA CCT TAT ATT ATT AAA AAA	546
	Lys Ile Ala Asp Asn Lys Pro Phe Val Lys Pro Tyr Ile Ile Lys Lys	
	140 145 150	
10	ATT GAT GGC GTG AGG GTG GCG GTT GTG GGG TAT GTG GTG GCG CAC ATC	594
	Ile Asp Gly Val Arg Val Ala Val Val Gly Tyr Val Val Ala His Ile	
	155 160 165	
	CCA ACT TGG GAG GCC TCT ACG CCT GAA CAT TTT GCA GGA TTG AAG TTT	642
15	Pro Thr Trp Glu Ala Ser Thr Pro Glu His Phe Ala Gly Leu Lys Phe	
	170 175 180	
	TTG GAC GCT GAA GAA GCG TTA AAA AAG ACC TTA AAA GAG TTG AAA GGG	690
	Leu Asp Ala Glu Glu Ala Leu Lys Lys Thr Leu Lys Glu Leu Lys Gly	
	185 190 195 200	
20	AAG TAT GAT ATT TTG ATT GGC GCT TTT CAT TTG GGG CGA GAA GAT GAG	738
	Lys Tyr Asp Ile Leu Ile Gly Ala Phe His Leu Gly Arg Glu Asp Glu	
	205 210 215	
	AAA GGT GGC GAC GGG ATA CCG GAT TTA GCG AAA AAA TTC CCG CAA TTT	786
	Lys Gly Gly Asp Gly Ile Pro Asp Leu Ala Lys Lys Phe Pro Gln Phe	
	220 225 230	
25	GAC ATC ATT TTT GCA GGG CAT GAG CAT GCG GTT TAT AAC ACC AAA GTA	834
	Asp Ile Ile Phe Ala Gly His Glu His Ala Val Tyr Asn Thr Lys Val	
	235 240 245	
	GGG AAA GTG CAT ACC ATT GAG CCT GGA GCG TAT GGG GCT TAT CTG GCA	882
30	Gly Lys Val His Thr Ile Glu Pro Gly Ala Tyr Gly Ala Tyr Leu Ala	
	250 255 260	
	AAG GGC GTG GTG GTA TTT GAC ACT AAA ACG AAG AAA AAA ATT ATA ACG	930
	Lys Gly Val Val Val Phe Asp Thr Lys Thr Lys Lys Lys Ile Ile Thr	
	265 270 275 280	
35	ACT GAA AAT TTA CCC ACA AAA GAT GTG CCA GAA GAT GAA GAA TTA GCG	978
	Thr Glu Asn Leu Pro Thr Lys Asp Val Pro Glu Asp Glu Glu Leu Ala	
	285 290 295	
	AAA AAA TAC GAA TAT GTG GAT AAA AAA TCA AAA GAA TAC GCT AAT GAA	1026
	Lys Lys Tyr Glu Tyr Val Asp Lys Lys Ser Lys Glu Tyr Ala Asn Glu	

	300	305	310	
	GTG GTT GGC GAA GTT ACA AAA ACC TTT ATT GAC AGG CCT GAT TTT ATC Val Val Gly Glu Val Thr Lys Thr Phe Ile Asp Arg Pro Asp Phe Ile 315 320 325	1074		
5	ACA GGA GAA GAA AAA ATC ACC ACG ATG CCC ACC GCC GCC TTG CAA GAA Thr Gly Glu Glu Lys Ile Thr Thr Met Pro Thr Ala Ala Leu Gln Glu 330 335 340	1122		
10	ACA CCG GTG ATA GAA TTG ATT AAT AAA GTG CAA AAA TAT TAC GCA AAA Thr Pro Val Ile Glu Leu Ile Asn Lys Val Gln Lys Tyr Tyr Ala Lys 345 350 355 360	1170		
	GCC GAT GTT TCA GCG GCA GCC TTA TTC AAT TTT GGG GCG AAT TTG AAA Ala Asp Val Ser Ala Ala Ala Leu Phe Asn Phe Gly Ala Asn Leu Lys 365 370 375	1218		
15	AAA GGG CCT TTC AAA AGA AAA GAT GTC ACT TAT ATT TAC AAG TTC GCT Lys Gly Pro Phe Lys Arg Lys Asp Val Thr Tyr Ile Tyr Lys Phe Ala 380 385 390	1266		
	AAT ACG CTC ATT GGA GTG CGT ATA ACG GGT GAA AAT CTG TTG AAA TAC Asn Thr Leu Ile Gly Val Arg Ile Thr Gly Glu Asn Leu Leu Lys Tyr 395 400 405	1314		
20	ATG GAA TGG TCA TAC CGA TTT TAC AAT CAG TTG CAA CCA GGA GAT TTG Met Glu Trp Ser Tyr Arg Phe Tyr Asn Gln Leu Gln Pro Gly Asp Leu 410 415 420	1362		
25	ACG ATC AGT TTT AAT GAA AAC ATT CGC GGC TAT AAC TTT GAT ATG TTT Thr Ile Ser Phe Asn Glu Asn Ile Arg Gly Tyr Asn Phe Asp Met Phe 425 430 435 440	1410		
	TCT GGC GTG AAA TAC CAG GTT GAT GTT ACA AAA CCC GCC GGA CAA AGG Ser Gly Val Lys Tyr Gln Val Asp Val Thr Lys Pro Ala Gly Gln Arg 445 450 455	1458		
30	ATT ATC AAT CCG ACA ATC AAC AAC AAA CCC ATT GAC CCC AAA GCC ATC Ile Ile Asn Pro Thr Ile Asn Asn Lys Pro Ile Asp Pro Lys Ala Ile 460 465 470	1506		
	TAT AAA TTA GCG ATC AAC AAT TAC CGA TTC GGA ACA TTA TCC ACG ACA Tyr Lys Leu Ala Ile Asn Asn Tyr Arg Phe Gly Thr Leu Ser Thr Thr 475 480 485	1554		
35	TTG AAT TTG GTT ACA GAC GCT GMT AGG TAT TAT AAT TCT TAC GAT GAA Leu Asn Leu Val Thr Asp Ala Xaa Arg Tyr Tyr Asn Ser Tyr Asp Glu 490 495 500	1602		

	Asn	Phe	Ser	Lys	Asp	Phe	Leu	Glu	Lys	Asn	Ile	Lys	Gly	Phe	Asn	Gly
		115						120					125			
	Asp	Val	Met	Asn	Ala	Asn	Ile	Ile	Lys	Ile	Ala	Asp	Asn	Lys	Pro	Phe
		130					135					140				
5	Val	Lys	Pro	Tyr	Ile	Ile	Lys	Lys	Ile	Asp	Gly	Val	Arg	Val	Ala	Val
	145					150					155					160
	Val	Gly	Tyr	Val	Val	Ala	His	Ile	Pro	Thr	Trp	Glu	Ala	Ser	Thr	Pro
				165						170					175	
10	Glu	His	Phe	Ala	Gly	Leu	Lys	Phe	Leu	Asp	Ala	Glu	Glu	Ala	Leu	Lys
				180					185					190		
	Lys	Thr	Leu	Lys	Glu	Leu	Lys	Gly	Lys	Tyr	Asp	Ile	Leu	Ile	Gly	Ala
			195					200					205			
	Phe	His	Leu	Gly	Arg	Glu	Asp	Glu	Lys	Gly	Gly	Asp	Gly	Ile	Pro	Asp
		210					215					220				
15	Leu	Ala	Lys	Lys	Phe	Pro	Gln	Phe	Asp	Ile	Ile	Phe	Ala	Gly	His	Glu
	225					230					235					240
	His	Ala	Val	Tyr	Asn	Thr	Lys	Val	Gly	Lys	Val	His	Thr	Ile	Glu	Pro
				245						250					255	
20	Gly	Ala	Tyr	Gly	Ala	Tyr	Leu	Ala	Lys	Gly	Val	Val	Val	Phe	Asp	Thr
				260					265					270		
	Lys	Thr	Lys	Lys	Lys	Ile	Ile	Thr	Thr	Glu	Asn	Leu	Pro	Thr	Lys	Asp
			275					280					285			
	Val	Pro	Glu	Asp	Glu	Glu	Leu	Ala	Lys	Lys	Tyr	Glu	Tyr	Val	Asp	Lys
		290					295				300					
25	Lys	Ser	Lys	Glu	Tyr	Ala	Asn	Glu	Val	Val	Gly	Glu	Val	Thr	Lys	Thr
	305					310					315					320
	Phe	Ile	Asp	Arg	Pro	Asp	Phe	Ile	Thr	Gly	Glu	Glu	Lys	Ile	Thr	Thr
				325						330					335	
30	Met	Pro	Thr	Ala	Ala	Leu	Gln	Glu	Thr	Pro	Val	Ile	Glu	Leu	Ile	Asn
				340					345				350			
	Lys	Val	Gln	Lys	Tyr	Tyr	Ala	Lys	Ala	Asp	Val	Ser	Ala	Ala	Ala	Leu
			355					360				365				
	Phe	Asn	Phe	Gly	Ala	Asn	Leu	Lys	Lys	Gly	Pro	Phe	Lys	Arg	Lys	Asp
		370				375					380					
35	Val	Thr	Tyr	Ile	Tyr	Lys	Phe	Ala	Asn	Thr	Leu	Ile	Gly	Val	Arg	Ile
	385					390					395					400
	Thr	Gly	Glu	Asn	Leu	Leu	Lys	Tyr	Met	Glu	Trp	Ser	Tyr	Arg	Phe	Tyr
				405						410					415	
40	Asn	Gln	Leu	Gln	Pro	Gly	Asp	Leu	Thr	Ile	Ser	Phe	Asn	Glu	Asn	Ile
				420					425					430		
	Arg	Gly	Tyr	Asn	Phe	Asp	Met	Phe	Ser	Gly	Val	Lys	Tyr	Gln	Val	Asp
		435						440				445				
	Val	Thr	Lys	Pro	Ala	Gly	Gln	Arg	Ile	Ile	Asn	Pro	Thr	Ile	Asn	Asn
		450					455				460					
45	Lys	Pro	Ile	Asp	Pro	Lys	Ala	Ile	Tyr	Lys	Leu	Ala	Ile	Asn	Asn	Tyr
	465					470					475					480
	Arg	Phe	Gly	Thr	Leu	Ser	Thr	Thr	Leu	Asn	Leu	Val	Thr	Asp	Ala	Xaa
				485						490					495	
50	Arg	Tyr	Tyr	Asn	Ser	Tyr	Asp	Glu	Leu	Gln	Asp	Asn	Gly	Gln	Ile	Arg
				500					505					510		

Asp Leu Ile Ile Lys Tyr Ile Thr Glu Glu Lys Gly Gly Lys Val Thr
 515 520 525
 Pro Glu Leu Glu Gly Asn Trp Glu Ile Ile Asn Tyr Asp Phe Lys Asn
 530 535 540
 5 Pro Leu Leu Glu Lys Leu Arg Glu Lys Leu Lys Glu Gly Ser Ile Lys
 545 550 555 560
 Ile Pro Thr Ser Lys Asp Gly Arg Thr Leu Asn Val Lys Ser Ile Lys
 565 570 575
 Glu Ser Glu Val Lys
 10 580

(2) INFORMATION FOR SEQ ID NO:73:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1339 base pairs
 (B) TYPE: nucleic acid
 15 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence
 (B) LOCATION: 68...1252
 20 (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:

	CCAATCGTTT AATAGCGATT AAATATGACT ATATACACTA CAACAATAAG ATTTTGAAAG	60
25	GTTGGTA ATG GAA TCA GTA AAA ACA GGA AAA ACA AAT AAG GTT GGC AAG	109
	Met Glu Ser Val Lys Thr Gly Lys Thr Asn Lys Val Gly Lys	
	1 5 10	
	AAT ACA GAG ATG GCT AAT ACA AAG GCA AAT AAA GAG GCT CAT TTT AAA	157
	Asn Thr Glu Met Ala Asn Thr Lys Ala Asn Lys Glu Ala His Phe Lys	
	15 20 25 30	
30	CAA GCG AGC ACC ATT ACA AAT ATA ATC AGA TCA ATT CGT GGG ATT TTT	205
	Gln Ala Ser Thr Ile Thr Asn Ile Ile Arg Ser Ile Arg Gly Ile Phe	
	35 40 45	
	ACA AAA ATT GCA AAG AAA GTT AGA GGA CTT GTA AAA AAA CAC CCC AAG	253
35	Thr Lys Ile Ala Lys Lys Val Arg Gly Leu Val Lys Lys His Pro Lys	
	50 55 60	
	AAA AGC AGT GCG GCA TTA GTA GTA TTG ACC CAT ATT GCG TGC AAG AAA	301
	Lys Ser Ser Ala Ala Leu Val Val Leu Thr His Ile Ala Cys Lys Lys	
	65 70 75	
	GCG AAA GAA TTA GAC GAT AAA GTC CAA GAT AAA TCC AAA CAA GCT GAA	349

	Ala	Lys	Glu	Leu	Asp	Asp	Lys	Val	Gln	Asp	Lys	Ser	Lys	Gln	Ala	Glu	
	80						85					90					
	AAA	GAA	AAT	CAA	ATC	AAT	TGG	TGG	AAA	TAT	TCA	GGA	TTA	ACA	ATA	GCG	397
5	Lys	Glu	Asn	Gln	Ile	Asn	Trp	Trp	Lys	Tyr	Ser	Gly	Leu	Thr	Ile	Ala	
	95					100				105					110		
	ACA	AGT	TTA	TTA	TTA	GCC	GCT	TGT	AGC	ACT	GGT	GAT	GTT	AGT	GAA	CAA	445
	Thr	Ser	Leu	Leu	Leu	Ala	Ala	Cys	Ser	Thr	Gly	Asp	Val	Ser	Glu	Gln	
						115				120					125		
10	ATA	GAA	CTA	GAA	CAA	GAA	AAA	CAA	AAG	ACG	AGC	AAT	ATA	GAG	ACT	AAC	493
	Ile	Glu	Leu	Glu	Gln	Glu	Lys	Gln	Lys	Thr	Ser	Asn	Ile	Glu	Thr	Asn	
				130					135					140			
	AAT	CAA	ATA	AAA	GTA	GAA	CAA	GAA	AAA	CAA	AAG	ACA	AGC	AAT	ATA	GAG	541
	Asn	Gln	Ile	Lys	Val	Glu	Gln	Glu	Lys	Gln	Lys	Thr	Ser	Asn	Ile	Glu	
				145					150					155			
15	ACT	AAT	AAT	CAA	ATA	AAA	GTA	GAA	CAA	GAA	CAA	CAA	AAG	ACA	AGC	AAT	589
	Thr	Asn	Asn	Gln	Ile	Lys	Val	Glu	Gln	Glu	Gln	Gln	Lys	Thr	Ser	Asn	
				160				165					170				
	ACA	CAG	AAA	GAT	TTG	GTT	AAA	GAA	CAG	AAA	GAT	TTG	GTT	AAA	GAA	CAG	637
20	Thr	Gln	Lys	Asp	Leu	Val	Lys	Glu	Gln	Lys	Asp	Leu	Val	Lys	Glu	Gln	
	175					180					185				190		
	AAA	GAT	TTG	GTT	AAA	GAA	CAG	AAA	GAT	TTG	GTT	AAA	GAA	CAG	AAA	GAT	685
	Lys	Asp	Leu	Val	Lys	Glu	Gln	Lys	Asp	Leu	Val	Lys	Glu	Gln	Lys	Asp	
						195				200				205			
25	TTG	GTT	AAA	ACA	CAG	AAA	GAT	TTC	ATT	AAA	TAT	GTA	GAA	CAA	AAT	TGC	733
	Leu	Val	Lys	Thr	Gln	Lys	Asp	Phe	Ile	Lys	Tyr	Val	Glu	Gln	Asn	Cys	
				210					215					220			
	CAA	GAA	AAT	CAT	AAT	CAA	TTC	TTT	ATT	GAA	AAA	GGA	GGA	ATT	AAG	GCT	781
	Gln	Glu	Asn	His	Asn	Gln	Phe	Phe	Ile	Glu	Lys	Gly	Gly	Ile	Lys	Ala	
				225				230					235				
30	GGT	ATT	GGT	ATA	GAA	GTA	GAA	GCT	GAA	TGC	AAA	ACC	CCT	AAA	CCT	GCA	829
	Gly	Ile	Gly	Ile	Glu	Val	Glu	Ala	Glu	Cys	Lys	Thr	Pro	Lys	Pro	Ala	
				240			245					250					
	AAA	ACC	AAT	CAA	ACC	CCT	ATC	CAG	CCA	AAA	CAC	CTC	CCA	AAC	TCT	AAA	877
35	Lys	Thr	Asn	Gln	Thr	Pro	Ile	Gln	Pro	Lys	His	Leu	Pro	Asn	Ser	Lys	
	255					260				265					270		
	CAA	CCC	CGC	TCT	CAA	AGA	GGA	TCA	AAA	GCG	CAA	GAG	CTT	ATC	GCT	TAT	925
	Gln	Pro	Arg	Ser	Gln	Arg	Gly	Ser	Lys	Ala	Gln	Glu	Leu	Ile	Ala	Tyr	
					275				280					285			

	TTG CAA AAA GAG CTA GAA TTT CTG CCC TAT TCG CAA AAA GCT ATC GCT	973
	Leu Gln Lys Glu Leu Glu Phe Leu Pro Tyr Ser Gln Lys Ala Ile Ala	
	290 295 300	
	AAA CAA GTG GAT TTT TAC AGG CCA AGT TCT ATC GCT TAT TTA GAA CTA	1021
5	Lys Gln Val Asp Phe Tyr Arg Pro Ser Ser Ile Ala Tyr Leu Glu Leu	
	305 310 315	
	GAT CCT AGA GAT TTT AAG GTT ACA GAA GAA TGG CAA AAA GAA AAT CTA	1069
	Asp Pro Arg Asp Phe Lys Val Thr Glu Glu Trp Gln Lys Glu Asn Leu	
	320 325 330	
10	AAA ATA CGC TCT AAA GCT CAA GCT AAA ATG CTT GAA ATG AGA AAC CCA	1117
	Lys Ile Arg Ser Lys Ala Gln Ala Lys Met Leu Glu Met Arg Asn Pro	
	335 340 345 350	
	CAA GCC CAC CTT TCA AAC TCT CAA AGC CTT TTG TTC GTT CAA AAA ATA	1165
	Gln Ala His Leu Ser Asn Ser Gln Ser Leu Leu Phe Val Gln Lys Ile	
15	355 360 365	
	TTT GCT GAT GTT AAT AAA GAA ATA GAA GCA GTT GCT AAT ACT GAA AAG	1213
	Phe Ala Asp Val Asn Lys Glu Ile Glu Ala Val Ala Asn Thr Glu Lys	
	370 375 380	
	AAA GCA GAA AAA GCG GGT TAT GGT TAT AGT AAA AGG ATG TAGCGGTTAA AA	1264
20	Lys Ala Glu Lys Ala Gly Tyr Gly Tyr Ser Lys Arg Met	
	385 390 395	
	ACATTGCACC AAGTTTTTAA TTATCTGTCG GCTTTTGAAA ACATTTTTTA TGGTAGCGTT	1324
	ATTTGGCAAT AAAAG	1339

(2) INFORMATION FOR SEQ ID NO:74:

25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 395 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

30 (ii) MOLECULE TYPE: protein
 (v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:

	Met Glu Ser Val Lys Thr Gly Lys Thr Asn Lys Val Gly Lys Asn Thr
	1 5 10 15
35	Glu Met Ala Asn Thr Lys Ala Asn Lys Glu Ala His Phe Lys Gln Ala
	20 25 30
	Ser Thr Ile Thr Asn Ile Ile Arg Ser Ile Arg Gly Ile Phe Thr Lys
	35 40 45

	Ile	Ala	Lys	Lys	Val	Arg	Gly	Leu	Val	Lys	Lys	His	Pro	Lys	Lys	Ser
	50						55					60				
	Ser	Ala	Ala	Leu	Val	Val	Leu	Thr	His	Ile	Ala	Cys	Lys	Lys	Ala	Lys
	65				70					75						80
5	Glu	Leu	Asp	Asp	Lys	Val	Gln	Asp	Lys	Ser	Lys	Gln	Ala	Glu	Lys	Glu
					85					90					95	
	Asn	Gln	Ile	Asn	Trp	Trp	Lys	Tyr	Ser	Gly	Leu	Thr	Ile	Ala	Thr	Ser
				100					105					110		
10	Leu	Leu	Leu	Ala	Ala	Cys	Ser	Thr	Gly	Asp	Val	Ser	Glu	Gln	Ile	Glu
			115					120					125			
	Leu	Glu	Gln	Glu	Lys	Gln	Lys	Thr	Ser	Asn	Ile	Glu	Thr	Asn	Asn	Gln
		130					135					140				
	Ile	Lys	Val	Glu	Gln	Glu	Lys	Gln	Lys	Thr	Ser	Asn	Ile	Glu	Thr	Asn
	145					150					155					160
15	Asn	Gln	Ile	Lys	Val	Glu	Gln	Glu	Gln	Gln	Lys	Thr	Ser	Asn	Thr	Gln
					165					170					175	
	Lys	Asp	Leu	Val	Lys	Glu	Gln	Lys	Asp	Leu	Val	Lys	Glu	Gln	Lys	Asp
				180					185					190		
20	Leu	Val	Lys	Glu	Gln	Lys	Asp	Leu	Val	Lys	Glu	Gln	Lys	Asp	Leu	Val
			195					200					205			
	Lys	Thr	Gln	Lys	Asp	Phe	Ile	Lys	Tyr	Val	Glu	Gln	Asn	Cys	Gln	Glu
		210					215					220				
	Asn	His	Asn	Gln	Phe	Phe	Ile	Glu	Lys	Gly	Gly	Ile	Lys	Ala	Gly	Ile
	225				230					235						240
25	Gly	Ile	Glu	Val	Glu	Ala	Glu	Cys	Lys	Thr	Pro	Lys	Pro	Ala	Lys	Thr
					245					250				255		
	Asn	Gln	Thr	Pro	Ile	Gln	Pro	Lys	His	Leu	Pro	Asn	Ser	Lys	Gln	Pro
				260					265					270		
30	Arg	Ser	Gln	Arg	Gly	Ser	Lys	Ala	Gln	Glu	Leu	Ile	Ala	Tyr	Leu	Gln
			275					280					285			
	Lys	Glu	Leu	Glu	Phe	Leu	Pro	Tyr	Ser	Gln	Lys	Ala	Ile	Ala	Lys	Gln
		290					295					300				
	Val	Asp	Phe	Tyr	Arg	Pro	Ser	Ser	Ile	Ala	Tyr	Leu	Glu	Leu	Asp	Pro
	305				310					315						320
35	Arg	Asp	Phe	Lys	Val	Thr	Glu	Glu	Trp	Gln	Lys	Glu	Asn	Leu	Lys	Ile
					325					330				335		
	Arg	Ser	Lys	Ala	Gln	Ala	Lys	Met	Leu	Glu	Met	Arg	Asn	Pro	Gln	Ala
				340					345				350			
40	His	Leu	Ser	Asn	Ser	Gln	Ser	Leu	Leu	Phe	Val	Gln	Lys	Ile	Phe	Ala
		355						360				365				
	Asp	Val	Asn	Lys	Glu	Ile	Glu	Ala	Val	Ala	Asn	Thr	Glu	Lys	Lys	Ala
		370					375					380				
	Glu	Lys	Ala	Gly	Tyr	Gly	Tyr	Ser	Lys	Arg	Met					
	385					390				395						

45 (2) INFORMATION FOR SEQ ID NO:75:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 904 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:

5 (A) NAME/KEY: Coding Sequence
(B) LOCATION: 70...864
(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:

	TAATAACTCA ATCCCATTTG AATGGCATTT TTAAGCCAAA TTGCTACTAT CTTTGGCTAA	60
10	AGGTTAAAC ATG ATT AAA CAA ACC CTC ATC ATT CTT GCC CCT TTT TTT ATC	111
	Met Ile Lys Gln Thr Leu Ile Ile Leu Ala Pro Phe Phe Ile	
	1 5 10	
	GCA ACG CTG TTG TAT TTT TTA GGC GCA CCG GAT GGG TTA AGA CCT AAC	159
	Ala Thr Leu Leu Tyr Phe Leu Gly Ala Pro Asp Gly Leu Arg Pro Asn	
15	15 20 25 30	
	GCT TGG CTT TAT TTT TGT ATT TTC ATG GGC ATG ATT ATA GGG CTA ATT	207
	Ala Trp Leu Tyr Phe Cys Ile Phe Met Gly Met Ile Ile Gly Leu Ile	
	35 40 45	
	TTA GAG CCG GTG CCA TCA GGT TTA ATA GCG CTA AGC GCG TTA GTG CTG	255
20	Leu Glu Pro Val Pro Ser Gly Leu Ile Ala Leu Ser Ala Leu Val Leu	
	50 55 60	
	TGT ATA GCG TTA AAA ATT GGA GCG AGC GAT AAA GTA GCG AGC GCT AAT	303
	Cys Ile Ala Leu Lys Ile Gly Ala Ser Asp Lys Val Ala Ser Ala Asn	
	65 70 75	
25	AAG GCT ATT TCG TGG GGT TTG AGC GGG TAT GCG AAT AAA ACG GTG TGG	351
	Lys Ala Ile Ser Trp Gly Leu Ser Gly Tyr Ala Asn Lys Thr Val Trp	
	80 85 90	
	CTT GTG TTT GTC GCT TTC ATT TTG GGT TTA GGG TAT GAA AAA AGC TTG	399
30	Leu Val Phe Val Ala Phe Ile Leu Gly Leu Gly Tyr Glu Lys Ser Leu	
	95 100 105 110	
	TTA GGG AAA CGG ATC GCT CTT TTA CTG ATT AGG TTT TTA GGG CAA ACC	447
	Leu Gly Lys Arg Ile Ala Leu Leu Leu Ile Arg Phe Leu Gly Gln Thr	
	115 120 125	
	CCT TTA GGT TTA GGC TAT GCG ATT GGT TTG AGC GAA TTG TGT CTA GCC	495
35	Pro Leu Gly Leu Gly Tyr Ala Ile Gly Leu Ser Glu Leu Cys Leu Ala	
	130 135 140	
	CCT TTT ATC CCT AGC AAC TCC GCT AGA AGT GGA GGC ATA CTC TAT CCC	543

Pro Phe Ile Pro Ser Asn Ser Ala Arg Ser Gly Gly Ile Leu Tyr Pro
 145 150 155

ATC GTT TCA TCT ATC CCG CCT TTA ATG GGA TCT ACT CCA AAT AAT AAC 591
 Ile Val Ser Ser Ile Pro Pro Leu Met Gly Ser Thr Pro Asn Asn Asn
 5 160 165 170

CCT GAC AAA ATC GGC GCG TAT TTG ATG TGG GTC GCT TTG GCT TCA ACT 639
 Pro Asp Lys Ile Gly Ala Tyr Leu Met Trp Val Ala Leu Ala Ser Thr
 175 180 185 190

10 TGC ATC ACT TCG TCC ATG TTT TTA ACC GCG CTC GCT CCT AAC CCC CTA 687
 Cys Ile Thr Ser Ser Met Phe Leu Thr Ala Leu Ala Pro Asn Pro Leu
 195 200 205

GCA ATG GAA ATC GCT GCC AAA ATG GGC GTG AAT GAA ATC TCA TGG TTT 735
 Ala Met Glu Ile Ala Ala Lys Met Gly Val Asn Glu Ile Ser Trp Phe
 210 215 220

15 TCG TGG TTT TTA GCG TTC TTG CCT TGT GGG GTG GTT TTG ATC TTG CTT 783
 Ser Trp Phe Leu Ala Phe Leu Pro Cys Gly Val Val Leu Ile Leu Leu
 225 230 235

20 GTG CCT TTA TTG GCG TAT AAA ACC TGC AAA CCC ACC TTA AAA GGC TCA 831
 Val Pro Leu Leu Ala Tyr Lys Thr Cys Lys Pro Thr Leu Lys Gly Ser
 240 245 250

AAA GAA GTG AGT TTG TGG GCC AAA AAA AGG AAT TAGAGGGCAT GGGGAGGTTT 884
 Lys Glu Val Ser Leu Trp Ala Lys Lys Arg Asn
 255 260 265

TCTTTAAAAG AAATTTTAAT 904

25 (2) INFORMATION FOR SEQ ID NO:76:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 265 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

30 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

35 Met Ile Lys Gln Thr Leu Ile Ile Leu Ala Pro Phe Phe Ile Ala Thr
 1 5 10 15
 Leu Leu Tyr Phe Leu Gly Ala Pro Asp Gly Leu Arg Pro Asn Ala Trp
 20 25 30

	Leu	Tyr	Phe	Cys	Ile	Phe	Met	Gly	Met	Ile	Ile	Gly	Leu	Ile	Leu	Glu
			35					40					45			
	Pro	Val	Pro	Ser	Gly	Leu	Ile	Ala	Leu	Ser	Ala	Leu	Val	Leu	Cys	Ile
		50					55					60				
5	Ala	Leu	Lys	Ile	Gly	Ala	Ser	Asp	Lys	Val	Ala	Ser	Ala	Asn	Lys	Ala
	65					70					75				80	
	Ile	Ser	Trp	Gly	Leu	Ser	Gly	Tyr	Ala	Asn	Lys	Thr	Val	Trp	Leu	Val
					85					90					95	
	Phe	Val	Ala	Phe	Ile	Leu	Gly	Leu	Gly	Tyr	Glu	Lys	Ser	Leu	Leu	Gly
10					100					105				110		
	Lys	Arg	Ile	Ala	Leu	Leu	Leu	Ile	Arg	Phe	Leu	Gly	Gln	Thr	Pro	Leu
			115					120					125			
	Gly	Leu	Gly	Tyr	Ala	Ile	Gly	Leu	Ser	Glu	Leu	Cys	Leu	Ala	Pro	Phe
	130						135					140				
15	Ile	Pro	Ser	Asn	Ser	Ala	Arg	Ser	Gly	Gly	Ile	Leu	Tyr	Pro	Ile	Val
	145					150					155				160	
	Ser	Ser	Ile	Pro	Pro	Leu	Met	Gly	Ser	Thr	Pro	Asn	Asn	Asn	Pro	Asp
					165					170					175	
	Lys	Ile	Gly	Ala	Tyr	Leu	Met	Trp	Val	Ala	Leu	Ala	Ser	Thr	Cys	Ile
20					180					185					190	
	Thr	Ser	Ser	Met	Phe	Leu	Thr	Ala	Leu	Ala	Pro	Asn	Pro	Leu	Ala	Met
			195					200					205			
	Glu	Ile	Ala	Ala	Lys	Met	Gly	Val	Asn	Glu	Ile	Ser	Trp	Phe	Ser	Trp
	210						215					220				
25	Phe	Leu	Ala	Phe	Leu	Pro	Cys	Gly	Val	Val	Leu	Ile	Leu	Leu	Val	Pro
	225					230					235				240	
	Leu	Leu	Ala	Tyr	Lys	Thr	Cys	Lys	Pro	Thr	Leu	Lys	Gly	Ser	Lys	Glu
				245						250					255	
	Val	Ser	Leu	Trp	Ala	Lys	Lys	Arg	Asn							
30				260					265							

(2) INFORMATION FOR SEQ ID NO:77:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1194 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 152...1069
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:

TTTTAAGCGG	TTTCCCTAAA	ATAGGTTTTT	AATCAATTTA	ATCCAAAGTT	GAATTTATTT	60
TTTGACAATA	TTATACTATA	ATAACCAATT	AGATTGGGGT	TTTACTGATT	TTTCTTTGTG	120

TGAGCTTTGG CTTAGTTTTG TAAGGAATGA G ATG ATA AAG AGT TGG ACT AAA 172
Met Ile Lys Ser Trp Thr Lys
1 5

5 AAG TGG TTT TTG ATT TTA TTT TTA ATG GCA AGT TGT TCC AGT TAT TTG 220
Lys Trp Phe Leu Ile Leu Phe Leu Met Ala Ser Cys Ser Ser Tyr Leu
10 15 20

GTG GCT ACA ACC GGT GAG AAA TAT TTT AAA ATG GCT ACT CAA GCC TTT 268
Val Ala Thr Thr Gly Glu Lys Tyr Phe Lys Met Ala Thr Gln Ala Phe
25 30 35

10 AAG AGA GGG GAC TAC CAT AAA GCG GTG GCT TTT TAT AAG AGG AGC TGT 316
Lys Arg Gly Asp Tyr His Lys Ala Val Ala Phe Tyr Lys Arg Ser Cys
40 45 50 55

AAT TTA AGG GTG GGG GTT GGT TGC ACG AGT TTA GGC TCT ATG TAT GAA 364
Asn Leu Arg Val Gly Val Gly Cys Thr Ser Leu Gly Ser Met Tyr Glu
15 60 65 70

GAT GGC GAT GGC GTG GAT CAG AAT ATT ACA AAA GCC GTT TTT TAT TAC 412
Asp Gly Asp Gly Val Asp Gln Asn Ile Thr Lys Ala Val Phe Tyr Tyr
75 80 85

20 AGA AGA GGG TGT AAT TTA AGG AAT CAT CTC GCT TGC GCG AGT CTA GGC 460
Arg Arg Gly Cys Asn Leu Arg Asn His Leu Ala Cys Ala Ser Leu Gly
90 95 100

TCT ATG TAT GAA GAT GGC GAT GGT GTG CAA AAA AAC CTT CCA AAG GCT 508
Ser Met Tyr Glu Asp Gly Asp Gly Val Gln Lys Asn Leu Pro Lys Ala
105 110 115

25 ATC TAT TAT TAC AGG AGA GGG TGC CAC TTA AAG GGT GGG GTG AGC TGT 556
Ile Tyr Tyr Tyr Arg Arg Gly Cys His Leu Lys Gly Gly Val Ser Cys
120 125 130 135

30 GGG AGT TTA GGT TTT ATG TAT TTT AAT GGC ACG GGC GTT AAG CAA AAT 604
Gly Ser Leu Gly Phe Met Tyr Phe Asn Gly Thr Gly Val Lys Gln Asn
140 145 150

TAT GCC AAA GCC CTT TTT CTT TCT AAA TAC GCT TGC AGT TTG AAT TAC 652
Tyr Ala Lys Ala Leu Phe Leu Ser Lys Tyr Ala Cys Ser Leu Asn Tyr
155 160 165

35 GGC ATT AGT TGT AAC TTT GTA GGG TAT ATG TAT AGG AAC GCC AAA GGC 700
Gly Ile Ser Cys Asn Phe Val Gly Tyr Met Tyr Arg Asn Ala Lys Gly
170 175 180

GTA CAG AAG GAT TTG AAA AAA GCC CTT GCG AAT TTT AAA AGA GGG TGC 748
Val Gln Lys Asp Leu Lys Lys Ala Leu Ala Asn Phe Lys Arg Gly Cys

185

190

195

CAT TTG AAA GAC GGA GCG AGT TGT GTG AGC TTG GGA TAC ATG TAT GAA 796
 His Leu Lys Asp Gly Ala Ser Cys Val Ser Leu Gly Tyr Met Tyr Glu
 200 205 210 215

5 GTC GGT ATG GAT GTC AAA CAA AAT GGA GAG CAA GCC TTG AAT CTT TAT 844
 Val Gly Met Asp Val Lys Gln Asn Gly Glu Gln Ala Leu Asn Leu Tyr
 220 225 230

10 AAA AAG GGT TGT TAT TTA AAA AGG GGG AGC GGT TGT CAT AAT GTG GCG 892
 Lys Lys Gly Cys Tyr Leu Lys Arg Gly Ser Gly Cys His Asn Val Ala
 235 240 245

GTG ATG TAT TAC ACC GGT AAG GGC GTT CCA AAG GAT TTA GAT AAA GCC 940
 Val Met Tyr Tyr Thr Gly Lys Gly Val Pro Lys Asp Leu Asp Lys Ala
 250 255 260

15 ATT TCG TAT TAT AAG AAA GGT TGC ACT CTA GGC TTT AGT GGT AGC TGT 988
 Ile Ser Tyr Tyr Lys Lys Gly Cys Thr Leu Gly Phe Ser Gly Ser Cys
 265 270 275

AAA GTG TTA GAA GAA GTG ATT GGC AAG AAG TCT GAT GAT TTG CAA GAT 1036
 Lys Val Leu Glu Glu Val Ile Gly Lys Lys Ser Asp Asp Leu Gln Asp
 280 285 290 295

20 GAC GCG CAA AAC GAC ACG CAA GAT GAT ATG CAA TAAGTTAAAG CTTATGGACT 1089
 Asp Ala Gln Asn Asp Thr Gln Asp Asp Met Gln
 300 305

AATGATTAAA ACTCATCTTA TAGAAATCTT TCTACTCTCT TGTTATCAAA TAGGGATTAA 1149
 GCGTCTCTAT TGATGGGTAT TGAGACTAAA AATCTGCAAA TCTAG 1194

25 (2) INFORMATION FOR SEQ ID NO:78:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 306 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:

35 Met Ile Lys Ser Trp Thr Lys Lys Trp Phe Leu Ile Leu Phe Leu Met
 1 5 10 15
 Ala Ser Cys Ser Ser Tyr Leu Val Ala Thr Thr Gly Glu Lys Tyr Phe
 20 25 30

	Lys	Met	Ala	Thr	Gln	Ala	Phe	Lys	Arg	Gly	Asp	Tyr	His	Lys	Ala	Val
		35						40					45			
	Ala	Phe	Tyr	Lys	Arg	Ser	Cys	Asn	Leu	Arg	Val	Gly	Val	Gly	Cys	Thr
	50						55					60				
5	Ser	Leu	Gly	Ser	Met	Tyr	Glu	Asp	Gly	Asp	Gly	Val	Asp	Gln	Asn	Ile
	65					70					75					80
	Thr	Lys	Ala	Val	Phe	Tyr	Tyr	Arg	Arg	Gly	Cys	Asn	Leu	Arg	Asn	His
				85						90					95	
	Leu	Ala	Cys	Ala	Ser	Leu	Gly	Ser	Met	Tyr	Glu	Asp	Gly	Asp	Gly	Val
10				100					105					110		
	Gln	Lys	Asn	Leu	Pro	Lys	Ala	Ile	Tyr	Tyr	Tyr	Arg	Arg	Gly	Cys	His
			115					120					125			
	Leu	Lys	Gly	Gly	Val	Ser	Cys	Gly	Ser	Leu	Gly	Phe	Met	Tyr	Phe	Asn
	130						135					140				
15	Gly	Thr	Gly	Val	Lys	Gln	Asn	Tyr	Ala	Lys	Ala	Leu	Phe	Leu	Ser	Lys
	145					150					155					160
	Tyr	Ala	Cys	Ser	Leu	Asn	Tyr	Gly	Ile	Ser	Cys	Asn	Phe	Val	Gly	Tyr
					165					170					175	
	Met	Tyr	Arg	Asn	Ala	Lys	Gly	Val	Gln	Lys	Asp	Leu	Lys	Lys	Ala	Leu
20				180					185					190		
	Ala	Asn	Phe	Lys	Arg	Gly	Cys	His	Leu	Lys	Asp	Gly	Ala	Ser	Cys	Val
			195					200					205			
	Ser	Leu	Gly	Tyr	Met	Tyr	Glu	Val	Gly	Met	Asp	Val	Lys	Gln	Asn	Gly
	210						215					220				
25	Glu	Gln	Ala	Leu	Asn	Leu	Tyr	Lys	Lys	Gly	Cys	Tyr	Leu	Lys	Arg	Gly
	225					230					235					240
	Ser	Gly	Cys	His	Asn	Val	Ala	Val	Met	Tyr	Tyr	Thr	Gly	Lys	Gly	Val
					245					250					255	
	Pro	Lys	Asp	Leu	Asp	Lys	Ala	Ile	Ser	Tyr	Tyr	Lys	Lys	Gly	Cys	Thr
30				260					265					270		
	Leu	Gly	Phe	Ser	Gly	Ser	Cys	Lys	Val	Leu	Glu	Glu	Val	Ile	Gly	Lys
			275					280					285			
	Lys	Ser	Asp	Asp	Leu	Gln	Asp	Asp	Ala	Gln	Asn	Asp	Thr	Gln	Asp	Asp
	290					295						300				
35	Met	Gln														
	305															

(2) INFORMATION FOR SEQ ID NO:79:

(i) SEQUENCE CHARACTERISTICS:

- 40 (A) LENGTH: 1001 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ix) FEATURE:

- 45 (A) NAME/KEY: Coding Sequence
 (B) LOCATION: 101...865

(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:

	TTTGTTTATA AGAAAAATTA TTTCAAATGT AGTAGAATTA AGGCAGTGTT TTTGCGTCAA	60
5	GCGATTTTAG GTTAATTTTG AGTTTTTAGG AGCAGTTTTT ATG CAA CAA GAA GAG	115
	Met Gln Gln Glu Glu	
	1 5	
	ATT ATA GAG GGT TAT TAT GGT GCT AGC AAA GGG CTT AAA AAG AGC GGT	163
10	Ile Ile Glu Gly Tyr Tyr Gly Ala Ser Lys Gly Leu Lys Lys Ser Gly	
	10 15 20	
	ATT TAT GCC AAG CTG GAT TTT TTA CAG AGC GCT ACG GGC TTG ATT TTA	211
	Ile Tyr Ala Lys Leu Asp Phe Leu Gln Ser Ala Thr Gly Leu Ile Leu	
	25 30 35	
	GCG CTC TTT ATG ATA GCA CAC ATG TTT TTA GTC TCA AGT ATC TTG ATT	259
15	Ala Leu Phe Met Ile Ala His Met Phe Leu Val Ser Ser Ile Leu Ile	
	40 45 50	
	AGC GAT GAA GCC ATG TAT AAA GTG GCG AAA TTT TTT GAA GGG AGC TTG	307
	Ser Asp Glu Ala Met Tyr Lys Val Ala Lys Phe Phe Glu Gly Ser Leu	
	55 60 65	
20	TTT TTA AAA GCG GGC GAG CCG GCT ATT GTG AGC GTG GTT GCA GCA GGG	355
	Phe Leu Lys Ala Gly Glu Pro Ala Ile Val Ser Val Val Ala Ala Gly	
	70 75 80 85	
	ATT ATT CTT ATT TTA GTC GCG CAT GCT TTT TTG GCG TTA AGG AAA TTC	403
25	Ile Ile Leu Ile Leu Val Ala His Ala Phe Leu Ala Leu Arg Lys Phe	
	90 95 100	
	CCT ATC AAT TAC AGG CAA TAC AAG GTT TTT AAA ACC CAT AAG CAT TTG	451
	Pro Ile Asn Tyr Arg Gln Tyr Lys Val Phe Lys Thr His Lys His Leu	
	105 110 115	
	ATG AAA CAT GGC GAT ACG AGC TTG TGG TTT ATT CAA GCC CTC ACC GGG	499
30	Met Lys His Gly Asp Thr Ser Leu Trp Phe Ile Gln Ala Leu Thr Gly	
	120 125 130	
	TTT GCG ATG TTT TTC TTA GCG AGT ATC CAC TTA TTT GTC ATG CTC ACA	547
	Phe Ala Met Phe Phe Leu Ala Ser Ile His Leu Phe Val Met Leu Thr	
	135 140 145	
35	GAG CCT GAA AGT ATT GGG CCT CAT GGT TCA AGC TAT CGT TTT GTA ACG	595
	Glu Pro Glu Ser Ile Gly Pro His Gly Ser Ser Tyr Arg Phe Val Thr	
	150 155 160 165	

CAA AAC TTT TGG CTT TTG TAT ATT TTC TTA TTG TTT GCC GTA GAA TTG 643
Gln Asn Phe Trp Leu Leu Tyr Ile Phe Leu Leu Phe Ala Val Glu Leu
170 175 180

CAT GGC TCT ATT GGG TTG TAT CGT TTA GCG ATC AAA TGG GGG TGG TTT 691
5 His Gly Ser Ile Gly Leu Tyr Arg Leu Ala Ile Lys Trp Gly Trp Phe
185 190 195

AAA AAT GTG AGC ATT CAA GGT TTG AGA AAA GTC AAA TGG GCG ATG AGC 739
Lys Asn Val Ser Ile Gln Gly Leu Arg Lys Val Lys Trp Ala Met Ser
200 205 210

10 GTG TTT TTT ATT GTT TTA GGG CTT TGC ACC TAT GGG GCT TAC ATT AAA 787
Val Phe Phe Ile Val Leu Gly Leu Cys Thr Tyr Gly Ala Tyr Ile Lys
215 220 225

AAA GGT TTA GAA AAT AAG GAA AAT GGC ATT AAA ACC ATG CAA GAA GCC 835
Lys Gly Leu Glu Asn Lys Glu Asn Gly Ile Lys Thr Met Gln Glu Ala
15 230 235 240 245

ATA GAA GCT GAT GGG AAA TTC CAC AAA GAA TAAGGGTAGA AAATGAAAAT AAC 888
Ile Glu Ala Asp Gly Lys Phe His Lys Glu
250 255

ATATTGTGAT GCGCTAATTA TTGGAGGCGG ACTAGCTGGG TTAAGGGCTA GTATCGCATG 948
20 CAAACAAAAG GGTTTAAACA CCATCGTTTT AAGCCTAGTG CCTGTCAGGC GTT 1001

(2) INFORMATION FOR SEQ ID NO:80:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 255 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:

30 Met Gln Gln Glu Glu Ile Ile Glu Gly Tyr Tyr Gly Ala Ser Lys Gly
1 5 10 15
Leu Lys Lys Ser Gly Ile Tyr Ala Lys Leu Asp Phe Leu Gln Ser Ala
20 25 30
Thr Gly Leu Ile Leu Ala Leu Phe Met Ile Ala His Met Phe Leu Val
35 40 45
Ser Ser Ile Leu Ile Ser Asp Glu Ala Met Tyr Lys Val Ala Lys Phe
50 55 60
Phe Glu Gly Ser Leu Phe Leu Lys Ala Gly Glu Pro Ala Ile Val Ser
65 70 75 80

	Val	Val	Ala	Ala	Gly	Ile	Ile	Leu	Ile	Leu	Val	Ala	His	Ala	Phe	Leu
					85					90					95	
	Ala	Leu	Arg	Lys	Phe	Pro	Ile	Asn	Tyr	Arg	Gln	Tyr	Lys	Val	Phe	Lys
				100					105					110		
5	Thr	His	Lys	His	Leu	Met	Lys	His	Gly	Asp	Thr	Ser	Leu	Trp	Phe	Ile
			115				120						125			
	Gln	Ala	Leu	Thr	Gly	Phe	Ala	Met	Phe	Phe	Leu	Ala	Ser	Ile	His	Leu
		130				135						140				
	Phe	Val	Met	Leu	Thr	Glu	Pro	Glu	Ser	Ile	Gly	Pro	His	Gly	Ser	Ser
10	145					150					155				160	
	Tyr	Arg	Phe	Val	Thr	Gln	Asn	Phe	Trp	Leu	Leu	Tyr	Ile	Phe	Leu	Leu
				165					170					175		
	Phe	Ala	Val	Glu	Leu	His	Gly	Ser	Ile	Gly	Leu	Tyr	Arg	Leu	Ala	Ile
				180					185					190		
15	Lys	Trp	Gly	Trp	Phe	Lys	Asn	Val	Ser	Ile	Gln	Gly	Leu	Arg	Lys	Val
			195				200						205			
	Lys	Trp	Ala	Met	Ser	Val	Phe	Phe	Ile	Val	Leu	Gly	Leu	Cys	Thr	Tyr
		210					215					220				
	Gly	Ala	Tyr	Ile	Lys	Lys	Gly	Leu	Glu	Asn	Lys	Glu	Asn	Gly	Ile	Lys
20	225					230					235				240	
	Thr	Met	Gln	Glu	Ala	Ile	Glu	Ala	Asp	Gly	Lys	Phe	His	Lys	Glu	
				245					250						255	

(2) INFORMATION FOR SEQ ID NO:81:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 975 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 82...912
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:

35	TTTTAAAATT AAAGAAAATT TTTTAAAG ATTATCACTC TTTTGTGATA AAGTAATCAT	60
	TTAAAATTTA GGGAGTTTTT T ATG GAA GAA TCA ACA GCG TTT ATT TTG GCT	111
	Met Glu Glu Ser Thr Ala Phe Ile Leu Ala	
	1 5 10	
	CTT GTG GGG CTA TTC ACC GGC ATT ACC GCC GGG TTT TTT GGT ATT GGT	159
40	Leu Val Gly Leu Phe Thr Gly Ile Thr Ala Gly Phe Phe Gly Ile Gly	
	15 20 25	
	GGG GGG GAG ATT GTC GTC CCT AGC GCG ATT TTT GCC CAT TTT AGC TAT	207

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

Gly Gly Glu Ile Val Val Pro Ser Ala Ile Phe Ala His Phe Ser Tyr
 30 35 40

	AGC	CAT	GCG	GTG	GGT	ATT	TCG	CTC	ATG	CAA	ATG	CTT	TTT	TCT	TCA	GTG	255
5	Ser	His	Ala	Val	Gly	Ile	Ser	Leu	Met	Gln	Met	Leu	Phe	Ser	Ser	Val	
			45					50					55				
	GTC	GGC	TCT	ATC	ATC	AAT	TAC	AAA	AAG	GGC	TTA	TTG	GAT	TTG	AGA	GAA	303
	Val	Gly	Ser	Ile	Ile	Asn	Tyr	Lys	Lys	Gly	Leu	Leu	Asp	Leu	Arg	Glu	
		60					65					70					
	GGC	TCA	TTT	GCC	GCG	CTT	GGA	GGG	CTA	ATG	GGA	GCG	ATT	TTA	GGG	AGC	351
10	Gly	Ser	Phe	Ala	Ala	Leu	Gly	Gly	Leu	Met	Gly	Ala	Ile	Leu	Gly	Ser	
	75					80				85					90		
	TTT	ATC	TTA	AAA	ATC	ATT	GAC	GAT	AAA	ATT	TTA	ATG	GCG	GTG	TTT	GTG	399
	Phe	Ile	Leu	Lys	Ile	Ile	Asp	Asp	Lys	Ile	Leu	Met	Ala	Val	Phe	Val	
					95					100					105		
15	GTG	GTG	GTG	TGC	TAC	ACC	TTT	ATC	AAA	TAC	GCT	TTT	TCT	AGC	AAC	AAG	447
	Val	Val	Val	Cys	Tyr	Thr	Phe	Ile	Lys	Tyr	Ala	Phe	Ser	Ser	Asn	Lys	
				110					115					120			
	AAA	CCC	AAG	CAT	TTT	GAA	GAA	ATG	CAT	TTT	GAT	TTG	CAT	GCG	AAT	AAC	495
20	Lys	Pro	Lys	His	Phe	Glu	Glu	Met	His	Phe	Asp	Leu	His	Ala	Asn	Asn	
			125					130					135				
	AAA	ACG	CCC	GAA	AAA	AAG	CGC	GCA	ATC	CCT	TTT	GTG	TCT	ATG	GAT	AGA	543
	Lys	Thr	Pro	Glu	Lys	Lys	Arg	Ala	Ile	Pro	Phe	Val	Ser	Met	Asp	Arg	
		140					145					150					
	ACG	CAT	GGG	GTT	TTG	ATG	CTC	GCC	GGT	TTT	GTT	ACC	GGC	ATC	TTT	TCT	591
25	Thr	His	Gly	Val	Leu	Met	Leu	Ala	Gly	Phe	Val	Thr	Gly	Ile	Phe	Ser	
	155					160					165				170		
	ATC	CCA	CTA	GGC	ATG	GGT	GGA	GGG	ATT	TTA	ATG	GTG	CCG	TTT	TTG	GGC	639
	Ile	Pro	Leu	Gly	Met	Gly	Gly	Gly	Ile	Leu	Met	Val	Pro	Phe	Leu	Gly	
				175					180					185			
30	TAT	TTT	TTG	AAA	TAC	GAT	TCT	AAA	AAA	ATC	GTG	CCT	TTG	GGG	CTA	TTT	687
	Tyr	Phe	Leu	Lys	Tyr	Asp	Ser	Lys	Lys	Ile	Val	Pro	Leu	Gly	Leu	Phe	
				190				195					200				
	TTT	GTG	GTG	TTC	GCT	TCT	TTA	TCT	GGG	GTC	ATC	TCT	CTT	TAT	AAC	GGG	735
35	Phe	Val	Val	Phe	Ala	Ser	Leu	Ser	Gly	Val	Ile	Ser	Leu	Tyr	Asn	Gly	
			205					210					215				
	AGG	GTT	CTT	GAT	AAT	ATA	AGC	GTT	CAA	GCG	GGG	GTG	ATT	ACC	GGC	ATT	783
	Arg	Val	Leu	Asp	Asn	Ile	Ser	Val	Gln	Ala	Gly	Val	Ile	Thr	Gly	Ile	
			220				225					230					

GGA GCG TTT TTA GGC GTG GGC ATT GGC ATC AAG CTT ATC GCT TTG GCT 831
 Gly Ala Phe Leu Gly Val Gly Ile Gly Ile Lys Leu Ile Ala Leu Ala
 235 240 245 250

AAT GAA AAG GTG CAT AAA ATC CTG TTG CTC CTT ATT TAT GCT TTA AGC 879
 5 Asn Glu Lys Val His Lys Ile Leu Leu Leu Leu Ile Tyr Ala Leu Ser
 255 260 265

ATT TTA GCG ACT TTA CAC AAG CTC ATT ATG GGG TAAATCTAAA AACGCTTCTA 932
 Ile Leu Ala Thr Leu His Lys Leu Ile Met Gly
 270 275

10 GGGCATT TTT AAAATTAATA TCAAAGAGCT TTCACCAGCA AGC 975

(2) INFORMATION FOR SEQ ID NO:82:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 277 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:82:

20 Met Glu Glu Ser Thr Ala Phe Ile Leu Ala Leu Val Gly Leu Phe Thr
 1 5 10 15
 Gly Ile Thr Ala Gly Phe Phe Gly Ile Gly Gly Gly Glu Ile Val Val
 20 25 30
 Pro Ser Ala Ile Phe Ala His Phe Ser Tyr Ser His Ala Val Gly Ile
 35 40 45
 25 Ser Leu Met Gln Met Leu Phe Ser Ser Val Val Gly Ser Ile Ile Asn
 50 55 60
 Tyr Lys Lys Gly Leu Leu Asp Leu Arg Glu Gly Ser Phe Ala Ala Leu
 65 70 75 80
 30 Gly Gly Leu Met Gly Ala Ile Leu Gly Ser Phe Ile Leu Lys Ile Ile
 85 90 95
 Asp Asp Lys Ile Leu Met Ala Val Phe Val Val Val Val Cys Tyr Thr
 100 105 110
 Phe Ile Lys Tyr Ala Phe Ser Ser Asn Lys Lys Pro Lys His Phe Glu
 115 120 125
 35 Glu Met His Phe Asp Leu His Ala Asn Asn Lys Thr Pro Glu Lys Lys
 130 135 140
 Arg Ala Ile Pro Phe Val Ser Met Asp Arg Thr His Gly Val Leu Met
 145 150 155 160
 40 Leu Ala Gly Phe Val Thr Gly Ile Phe Ser Ile Pro Leu Gly Met Gly
 165 170 175
 Gly Gly Ile Leu Met Val Pro Phe Leu Gly Tyr Phe Leu Lys Tyr Asp

		180		185		190
	Ser	Lys	Lys	Ile	Val	Pro
		195		200		205
	Leu	Ser	Gly	Val	Ile	Ser
5		210		215		220
	Ser	Val	Gln	Ala	Gly	Val
	225		230		235	240
	Gly	Ile	Gly	Ile	Lys	Leu
		245		250		255
10	Ile	Leu	Leu	Leu	Ile	Tyr
		260		265		270
	Lys	Leu	Ile	Met	Gly	
		275				

(2) INFORMATION FOR SEQ ID NO:83:

15 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1667 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

20 (ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 220...1482
- (D) OTHER INFORMATION:

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:83:

AAGCGCGAGC	TATATGAGGA	ATTTTAGCTT	CTATGTGGGC	TATTCAGTCG	GTTTTTAAGG	60
AAGGCTCTTG	ATGAAAAATA	CCAATACAAA	AGAGATAAAG	AATACAAGAA	TGAAAAAAGG	120
TTATAGTCAA	TACCACGCGC	TCAAAAAAGG	GCTTTTAAAA	ACGCTCTGCT	TTTGTAGCCTT	180
CCTTTAAGCG	TGGCGTTAGC	TGAAGACGAT	GGCTTTTAT	ATG GGA GTG GGC TAT		234
				Met Gly Val Gly Tyr		
				1 5		

CAA ATC GGC GGC GCG CAA CAA AAT ATC GAT AAC AAA GGC AGC ACC CTA	282
Gln Ile Gly Gly Ala Gln Gln Asn Ile Asp Asn Lys Gly Ser Thr Leu	
10 15 20	

35 AGG AAT AAT GTC ATT AAT AAT TTC CGC CAA GTG GGC GTG GGT ATG GCA	330
Arg Asn Asn Val Ile Asn Asn Phe Arg Gln Val Gly Val Gly Met Ala	
25 30 35	

GGG GGT AAT GGG CTT TTA GCC TTA GCG ACA AAC ACG ACC ATG GAC GCT	378
Gly Gly Asn Gly Leu Leu Ala Leu Ala Thr Asn Thr Thr Met Asp Ala	
40 45 50	

	CTT TTA GGG ATA GGC AAC CAA ATT GTC AAT ACT AAT ACA ACT GTT AGC	426
	Leu Leu Gly Ile Gly Asn Gln Ile Val Asn Thr Asn Thr Thr Val Ser	
	55 60 65	
5	AAC AAC AAC GCA GAA TTA ACC CAG TTT AAA AAA ATA CTC CCT CAA ATT	474
	Asn Asn Asn Ala Glu Leu Thr Gln Phe Lys Lys Ile Leu Pro Gln Ile	
	70 75 80 85	
	GAG CAA CGC TTT GAA ACG AAT AAA AAC GCT TAT AGC GTT CAA GCC TTG	522
	Glu Gln Arg Phe Glu Thr Asn Lys Asn Ala Tyr Ser Val Gln Ala Leu	
	90 95 100	
10	CAA GTG TAT TTG AGT AAT GTG CTT TAT AAC TTG GTT AAT AAT AGT AAT	570
	Gln Val Tyr Leu Ser Asn Val Leu Tyr Asn Leu Val Asn Asn Ser Asn	
	105 110 115	
	AAT GGC AGT AAT AAT GGA GTC GTT CCT GAA TAT GTA GGA ATT ATA AAA	618
	Asn Gly Ser Asn Asn Gly Val Val Pro Glu Tyr Val Gly Ile Ile Lys	
15	120 125 130	
	GTT CTC TAT GGT TCT CAA AAT GAA TTC AGT CTC TTA GCC ACG GAG AGT	666
	Val Leu Tyr Gly Ser Gln Asn Glu Phe Ser Leu Leu Ala Thr Glu Ser	
	135 140 145	
	GTG GTG CTT TTA AAC GCG CTT ACA AGG GTG AAT CTG GAT AGT AAT TCG	714
20	Val Val Leu Leu Asn Ala Leu Thr Arg Val Asn Leu Asp Ser Asn Ser	
	150 155 160 165	
	GTG TTT TTA AAA GGG CTA TTA GCC CAA ATG CAG CTT TTT AAT GAC ACT	762
	Val Phe Leu Lys Gly Leu Leu Ala Gln Met Gln Leu Phe Asn Asp Thr	
	170 175 180	
25	TCT TCA GCA AAG CTA GGC CAG ATC GCA GAA AAC TTG AAG AAC GGT GGT	810
	Ser Ser Ala Lys Leu Gly Gln Ile Ala Glu Asn Leu Lys Asn Gly Gly	
	185 190 195	
	GCA GGA TCA ATG CTC CAA AAG GAT GTG AAA ACC ATC TCG GAT CGA ATC	858
30	Ala Gly Ser Met Leu Gln Lys Asp Val Lys Thr Ile Ser Asp Arg Ile	
	200 205 210	
	GCT ACT TAC CAA GAG AAT CTA AAA CAG CTA GGA GGG ATG CTA AAG AAT	906
	Ala Thr Tyr Gln Glu Asn Leu Lys Gln Leu Gly Gly Met Leu Lys Asn	
	215 220 225	
	TAC GAT GAA CCC TAC TTG CCC CAA TTT GGG CCA GGC ACA AGC TCT CAG	954
35	Tyr Asp Glu Pro Tyr Leu Pro Gln Phe Gly Pro Gly Thr Ser Ser Gln	
	230 235 240 245	
	CAT GGG GTT ATT AAT GGC TTT GGC ATT CAA GTG GGC TAT AAG CAA TTT	1002
	His Gly Val Ile Asn Gly Phe Gly Ile Gln Val Gly Tyr Lys Gln Phe	

	250	255	260	
	TTT GGG AAC AAG CGG AAT ATA GGC TTA CGA TAT TAC GCT TTC TTT GAT			1050
	Phe Gly Asn Lys Arg Asn Ile Gly Leu Arg Tyr Tyr Ala Phe Phe Asp			
	265	270	275	
5	TAT GGC TTT ACG CAA TTG GGC AGT CTT AGC AGC GCC GTT AAA GCG AAT			1098
	Tyr Gly Phe Thr Gln Leu Gly Ser Leu Ser Ser Ala Val Lys Ala Asn			
	280	285	290	
	ATC TTT ACT TAT GGC GCT GGC ACG GAC TTT TTA TGG AAT ATC TTT AGA			1146
10	Ile Phe Thr Tyr Gly Ala Gly Thr Asp Phe Leu Trp Asn Ile Phe Arg			
	295	300	305	
	AGG GTT TTT AGC GAT CAG TCC TTG AAT GTG GGG GTG TTT GGG GGC ATT			1194
	Arg Val Phe Ser Asp Gln Ser Leu Asn Val Gly Val Phe Gly Gly Ile			
	310	315	320	325
	CAA ATA GCG GGT AAC ACT TGG GAT AGC TCT TTA AGA GGT CAA ATT GAA			1242
15	Gln Ile Ala Gly Asn Thr Trp Asp Ser Ser Leu Arg Gly Gln Ile Glu			
	330	335	340	
	AAC TCG TTT AAA GAA TAC CCC ACT CCC ACG AAT TTC CAA TTT TTG TTT			1290
	Asn Ser Phe Lys Glu Tyr Pro Thr Pro Thr Asn Phe Gln Phe Leu Phe			
	345	350	355	
20	AAT TTG GGT TTA AGG GCT CAT TTT GCC AGC ACC ATG CAC CGC CGG TTT			1338
	Asn Leu Gly Leu Arg Ala His Phe Ala Ser Thr Met His Arg Arg Phe			
	360	365	370	
	TTG AGC GCG TCT CAA AGC ATT CAG CAT GGG ATG GAA TTT GGC GTG AAA			1386
25	Leu Ser Ala Ser Gln Ser Ile Gln His Gly Met Glu Phe Gly Val Lys			
	375	380	385	
	ATC CCG GCT ATC AAT CAA AGG TAT TTG AGG GCC AAT GGG GCT GAT GTG			1434
	Ile Pro Ala Ile Asn Gln Arg Tyr Leu Arg Ala Asn Gly Ala Asp Val			
	390	395	400	405
	GAT TAC AGG CGT TTG TAT GCG TTC TAT ATC AAT TAC ACG ATA GGT TTT T			1483
30	Asp Tyr Arg Arg Leu Tyr Ala Phe Tyr Ile Asn Tyr Thr Ile Gly Phe			
	410	415	420	
	AAGCTCTTTT TAGGGCTTAT AAAGAGGCTT TTTACTTTTT TTTTGGTATT CTAACAAGCT			1543
	TTTAAATAAT CCAATCTACT TTGTTTAAAG GATAATATTT TATGGCAGAT GTCGTTGTGG			1603
	GGATCCAGTG GGGAGATGAG GGGGAAGGGAA AAATTGTTGA TAGGATCGCT AAAGATTATG			1663
35	ACTT			1667

(2) INFORMATION FOR SEQ ID NO:84:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 421 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

5 (ii) MOLECULE TYPE: protein
 (v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:84:

	Met	Gly	Val	Gly	Tyr	Gln	Ile	Gly	Gly	Ala	Gln	Gln	Asn	Ile	Asp	Asn	
	1				5					10					15		
10	Lys	Gly	Ser	Thr	Leu	Arg	Asn	Asn	Val	Ile	Asn	Asn	Phe	Arg	Gln	Val	
				20					25				30				
	Gly	Val	Gly	Met	Ala	Gly	Gly	Asn	Gly	Leu	Leu	Ala	Leu	Ala	Thr	Asn	
				35				40					45				
	Thr	Thr	Met	Asp	Ala	Leu	Leu	Gly	Ile	Gly	Asn	Gln	Ile	Val	Asn	Thr	
15		50					55					60					
	Asn	Thr	Thr	Val	Ser	Asn	Asn	Asn	Ala	Glu	Leu	Thr	Gln	Phe	Lys	Lys	
	65					70					75				80		
	Ile	Leu	Pro	Gln	Ile	Glu	Gln	Arg	Phe	Glu	Thr	Asn	Lys	Asn	Ala	Tyr	
				85					90						95		
20	Ser	Val	Gln	Ala	Leu	Gln	Val	Tyr	Leu	Ser	Asn	Val	Leu	Tyr	Asn	Leu	
				100					105					110			
	Val	Asn	Asn	Ser	Asn	Asn	Gly	Ser	Asn	Asn	Gly	Val	Val	Pro	Glu	Tyr	
				115				120					125				
	Val	Gly	Ile	Ile	Lys	Val	Leu	Tyr	Gly	Ser	Gln	Asn	Glu	Phe	Ser	Leu	
25		130					135					140					
	Leu	Ala	Thr	Glu	Ser	Val	Val	Leu	Leu	Asn	Ala	Leu	Thr	Arg	Val	Asn	
	145					150					155				160		
	Leu	Asp	Ser	Asn	Ser	Val	Phe	Leu	Lys	Gly	Leu	Leu	Ala	Gln	Met	Gln	
				165					170						175		
30	Leu	Phe	Asn	Asp	Thr	Ser	Ser	Ala	Lys	Leu	Gly	Gln	Ile	Ala	Glu	Asn	
				180					185					190			
	Leu	Lys	Asn	Gly	Gly	Ala	Gly	Ser	Met	Leu	Gln	Lys	Asp	Val	Lys	Thr	
			195				200						205				
	Ile	Ser	Asp	Arg	Ile	Ala	Thr	Tyr	Gln	Glu	Asn	Leu	Lys	Gln	Leu	Gly	
35		210					215					220					
	Gly	Met	Leu	Lys	Asn	Tyr	Asp	Glu	Pro	Tyr	Leu	Pro	Gln	Phe	Gly	Pro	
	225					230					235				240		
	Gly	Thr	Ser	Ser	Gln	His	Gly	Val	Ile	Asn	Gly	Phe	Gly	Ile	Gln	Val	
				245						250					255		
40	Gly	Tyr	Lys	Gln	Phe	Phe	Gly	Asn	Lys	Arg	Asn	Ile	Gly	Leu	Arg	Tyr	
				260					265					270			
	Tyr	Ala	Phe	Phe	Asp	Tyr	Gly	Phe	Thr	Gln	Leu	Gly	Ser	Leu	Ser	Ser	
			275					280					285				
	Ala	Val	Lys	Ala	Asn	Ile	Phe	Thr	Tyr	Gly	Ala	Gly	Thr	Asp	Phe	Leu	
45		290					295					300					
	Trp	Asn	Ile	Phe	Arg	Arg	Val	Phe	Ser	Asp	Gln	Ser	Leu	Asn	Val	Gly	
	305					310					315				320		

	Val	Phe	Gly	Gly	Ile	Gln	Ile	Ala	Gly	Asn	Thr	Trp	Asp	Ser	Ser	Leu	
					325					330						335	
	Arg	Gly	Gln	Ile	Glu	Asn	Ser	Phe	Lys	Glu	Tyr	Pro	Thr	Pro	Thr	Asn	
				340					345					350			
5	Phe	Gln	Phe	Leu	Phe	Asn	Leu	Gly	Leu	Arg	Ala	His	Phe	Ala	Ser	Thr	
			355				360						365				
	Met	His	Arg	Arg	Phe	Leu	Ser	Ala	Ser	Gln	Ser	Ile	Gln	His	Gly	Met	
		370				375						380					
	Glu	Phe	Gly	Val	Lys	Ile	Pro	Ala	Ile	Asn	Gln	Arg	Tyr	Leu	Arg	Ala	
10	385					390					395					400	
	Asn	Gly	Ala	Asp	Val	Asp	Tyr	Arg	Arg	Leu	Tyr	Ala	Phe	Tyr	Ile	Asn	
				405					410						415		
	Tyr	Thr	Ile	Gly	Phe												
				420													

15 (2) INFORMATION FOR SEQ ID NO:85:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 926 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 207...746
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:85:

	CCCCCTTAATT	GCAGATGTTT	TGCAAGAGGG	ATTGCGTGGC	GTCTATCATT	CTAGAGAGAT	60
	AGACTTTGTA	GAAAAAGTGG	TTGTTT TAGA	CAGCTGTCAA	ATCCACCAAA	AAGCGTTAAT	120
	GCATTTGCAA	GAAACTTTGA	TGATAGAAGT	GGATAGGCTT	GATTTTCTT	TAGTGGAGCG	180
30	CTTGAACATT	TTAGCGCGCA	TGGAGA	ATG AAA AGC	ATG CGT TTT	AGT TAC ATT	233
				Met Lys Ser Met	Arg Phe Ser Tyr	Ile	
				1		5	
	GAG CCA AGA	GCG AAA TAC	CTT ATC AGC	AAG CTT TCT	AAA ATT TGG	GTT	281
	Glu Pro Arg	Ala Lys Tyr	Leu Ile Ser	Lys Leu Ser	Lys Ile Trp	Val	
35	10		15		20	25	
	TTT TAC ATT	TTT TTA TCT	TTT GTG	GTA ATA GGG	GGG TTA GTG	TGG TTT	329
	Phe Tyr Ile	Phe Leu Ser	Phe Val Val	Ile Gly Gly	Leu Val Trp	Phe	
		30		35		40	
	ATG CAC AAC	GCC ATT AAA	AGC ACT CAA	GAC AAC GCG	TCC AGT TTG	ACG	377
40	Met His Asn	Ala Ile Lys	Ser Thr Gln	Asp Asn Ala	Ser Ser Leu	Thr	
		45		50		55	

	ATC CAA GAA AGG CTC TAC CGC CAT GAA ATC AGC CGC TTA CAG GTT AAG	425
	Ile Gln Glu Arg Leu Tyr Arg His Glu Ile Ser Arg Leu Gln Val Lys	
	60 65 70	
	ACT GAT GAA ACC TTA AAA CTC ATT AAA GAA GCC AAA AAG CGT TTG AAT	473
5	Thr Asp Glu Thr Leu Lys Leu Ile Lys Glu Ala Lys Lys Arg Leu Asn	
	75 80 85	
	TAT AAC GAT GAT ATA CGA GAT GTT TTG CAA GGG CTT TTG AAT ATT GTG	521
	Tyr Asn Asp Asp Ile Arg Asp Val Leu Gln Gly Leu Leu Asn Ile Val	
	90 95 100 105	
10	CCG GAT TCC ATC ACT ATT AAT AGC ATT GAA ATA GAC CAG CAA AGC GTG	569
	Pro Asp Ser Ile Thr Ile Asn Ser Ile Glu Ile Asp Gln Gln Ser Val	
	110 115 120	
	GTT GTT AGC GGT AAA ACC CCT TCT AAA GAA GCC TTT TAT TTT TTG TTT	617
	Val Val Ser Gly Lys Thr Pro Ser Lys Glu Ala Phe Tyr Phe Leu Phe	
15	125 130 135	
	CAA AAC AAA CTA AAC CCC ATG TTT GAT TAT TCT AGG GCG GAA TTT TTC	665
	Gln Asn Lys Leu Asn Pro Met Phe Asp Tyr Ser Arg Ala Glu Phe Phe	
	140 145 150	
	CCC TTA AGC GAT GGG TGG TTT AAT TTT GTC TCC ACC AAC TTT TCT AAT	713
20	Pro Leu Ser Asp Gly Trp Phe Asn Phe Val Ser Thr Asn Phe Ser Asn	
	155 160 165	
	TCC TTA CTG ATA AAA AAT CCG GAG TCT ATT AAA TGAAGCCATT GCATTTTTC	766
	Ser Leu Leu Ile Lys Asn Pro Glu Ser Ile Lys	
	170 175 180	
25	CACCTGGACA GAGAGCAATC AGGCGATGTG GGGTTTATCA TTAAAAACCT CGTTTTTTTA	826
	GGGGTTTTTT CCTTATTGGG TTGGTTGAAT ACCGAGTATT TTCTATGGCC TAGCATGCTG	886
	GAATTAAAAA AAATCCTTTT AGAAGAAAAT CGTAAAAAAA	926

(2) INFORMATION FOR SEQ ID NO:86:

30 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 180 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

35 (ii) MOLECULE TYPE: protein
 (v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:86:

Met Lys Ser Met Arg Phe Ser Tyr Ile Glu Pro Arg Ala Lys Tyr Leu

	AAG CTT ATT TTG ATT GAT GAA TTT GAA GTT TTA GCC AAT AAA TTC ATT	270
	Lys Leu Ile Leu Ile Asp Glu Phe Glu Val Leu Ala Asn Lys Phe Ile	
	25 30 35 40	
5	TCT CGT TTG CCC AAT ATC CCT AGC ACC CCT AGA GAG TTT GGG TTA GGC	318
	Ser Arg Leu Pro Asn Ile Pro Ser Thr Pro Arg Glu Phe Gly Leu Gly	
	45 50 55	
	AAG GGC GAG ATC ATG GAG ATT GAT GTG CCT TTT GGG AGT ATT TTT GCT	366
	Lys Gly Glu Ile Met Glu Ile Asp Val Pro Phe Gly Ser Ile Phe Ala	
	60 65 70	
10	TAC AGA CAC ATT GGC TCT ATC AGA CAA AAA GAA TAC AGG ATT GTA GGG	414
	Tyr Arg His Ile Gly Ser Ile Arg Gln Lys Glu Tyr Arg Ile Val Gly	
	75 80 85	
15	CTT TAT CGC AAC GAT GTT TTG TTG CTC TCC ACT AAA TCT TTA GTT ATC	462
	Leu Tyr Arg Asn Asp Val Leu Leu Leu Ser Thr Lys Ser Leu Val Ile	
	90 95 100	
	CAG CCG CGA GAC ATT CTC TTA GTG GCG GGT AAT CCG GAA ATT TTG AAT	510
	Gln Pro Arg Asp Ile Leu Leu Val Ala Gly Asn Pro Glu Ile Leu Asn	
	105 110 115 120	
20	GCG GTG TAT CTT CAA GTC AAA AGC AAT GTG GGG CAG TTC CCA GCC CCC	558
	Ala Val Tyr Leu Gln Val Lys Ser Asn Val Gly Gln Phe Pro Ala Pro	
	125 130 135	
	TTT GGT AAG AGC ATT TAT TTA TAC ATT GAT ATG CGT TTG CAG AAC AGA	606
	Phe Gly Lys Ser Ile Tyr Leu Tyr Ile Asp Met Arg Leu Gln Asn Arg	
	140 145 150	
25	AAA GCG ATG ATG CGC GAT GTG TAT CAA GCC TTG TTT TTG CAC AAA CAT	654
	Lys Ala Met Met Arg Asp Val Tyr Gln Ala Leu Phe Leu His Lys His	
	155 160 165	
30	TTA AAG AGC TAC AAG CTC TAC ATT CAG GTT TTA CAC CCC ACT AGC CCT	702
	Leu Lys Ser Tyr Lys Leu Tyr Ile Gln Val Leu His Pro Thr Ser Pro	
	170 175 180	
	AAG TTT TAC CAT AAA TTT TTA GCG CTA GAA ACC GAA AGC ATT GAA GTG	750
	Lys Phe Tyr His Lys Phe Leu Ala Leu Glu Thr Glu Ser Ile Glu Val	
	185 190 195 200	
35	AAT TTT GAT TTT TAC AGG AAA AGT TTT ATC CAA AAA CTC CAT GAA GAC	798
	Asn Phe Asp Phe Tyr Arg Lys Ser Phe Ile Gln Lys Leu His Glu Asp	
	205 210 215	
	CAC CAG AAA AAA ATG GGC CTA ATC GTG GTA GGC AGA GAG CTT TTT TTA	846
	His Gln Lys Lys Met Gly Leu Ile Val Val Gly Arg Glu Leu Phe Leu	

	220	225	230	
	TCT AAA AAA CAC CGA AAG GCC TTG TAT AAA ACA GCC ACC CCA GTT TAT			894
	Ser Lys Lys His Arg Lys Ala Leu Tyr Lys Thr Ala Thr Pro Val Tyr			
	235	240	245	
5	AAA ACC AAC ACT TCT GGC TTG TCT AAA ACC TCT CAA AGC GTG GTG GTA			942
	Lys Thr Asn Thr Ser Gly Leu Ser Lys Thr Ser Gln Ser Val Val Val			
	250	255	260	
	TTG AAT GAA AGT TTG GAT ATT AAT GAG GAC ATG TCT TCA GTG ATT TTT			990
	Leu Asn Glu Ser Leu Asp Ile Asn Glu Asp Met Ser Ser Val Ile Phe			
10	265	270	275	280
	GAT GTG TCT ATG CAA ATG GAT TTG GGC TTG TTG CTC TAT GAT TTT GAC			1038
	Asp Val Ser Met Gln Met Asp Leu Gly Leu Leu Leu Tyr Asp Phe Asp			
	285	290	295	
	CCT AAC AAG CGC TAT AAA AAC GAG ATT GTC AAT CAT TAT GAA AAT TTA			1086
15	Pro Asn Lys Arg Tyr Lys Asn Glu Ile Val Asn His Tyr Glu Asn Leu			
	300	305	310	
	GCC AAC GCG TTC AAC CGC AAG ATT GAG ATT TTC CAA ACC GAT ATT AGA			1134
	Ala Asn Ala Phe Asn Arg Lys Ile Glu Ile Phe Gln Thr Asp Ile Arg			
	315	320	325	
20	AAT CCT ATC ATG TAT CTC AAT TCT TTA AGA AAT CCC ATT TTG CAT TTC			1182
	Asn Pro Ile Met Tyr Leu Asn Ser Leu Arg Asn Pro Ile Leu His Phe			
	330	335	340	
	ATG CCT TTT GAA GAG TGC ATC ACG CAC ACG CGC TTT TGG TGG TTT TTA			1230
	Met Pro Phe Glu Glu Cys Ile Thr His Thr Arg Phe Trp Trp Phe Leu			
25	345	350	355	360
	TCC ACT AAA GTG GAA AAA TTA GCG TTT TTA AAC GAT GAT AAC CCT CAA			1278
	Ser Thr Lys Val Glu Lys Leu Ala Phe Leu Asn Asp Asp Asn Pro Gln			
	365	370	375	
	ATT TTT ATC CCT GTA GCG GAG TGAAAGAATG CAAGAAATTT TAATCCCTTT AAAA			1333
30	Ile Phe Ile Pro Val Ala Glu			
	380			
	GAAAAAACT ATAAAGTGTT TTTGGGGGAA CTGCCTGAAA TAAAATTGAA ACAAAAAGCC			1393
	CTCATCATTA GCGATAGCAT CGTAGCCGGG TTGCATTGTC CCTATTT			1440

(2) INFORMATION FOR SEQ ID NO:88:

- 35 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 383 amino acids
 (B) TYPE: amino acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:88:

	Met	Cys	Val	Val	Leu	Ser	Val	Lys	Arg	Asp	Gly	Glu	Lys	Thr	Leu	Glu
	1				5				10					15		
	Asn	Asn	Glu	Glu	Asn	Lys	Asp	Glu	Lys	Leu	Ile	Leu	Ile	Asp	Glu	Phe
			20					25					30			
10	Glu	Val	Leu	Ala	Asn	Lys	Phe	Ile	Ser	Arg	Leu	Pro	Asn	Ile	Pro	Ser
		35					40				45					
	Thr	Pro	Arg	Glu	Phe	Gly	Leu	Gly	Lys	Gly	Glu	Ile	Met	Glu	Ile	Asp
	50				55				60							
	Val	Pro	Phe	Gly	Ser	Ile	Phe	Ala	Tyr	Arg	His	Ile	Gly	Ser	Ile	Arg
15	65				70				75					80		
	Gln	Lys	Glu	Tyr	Arg	Ile	Val	Gly	Leu	Tyr	Arg	Asn	Asp	Val	Leu	Leu
				85				90					95			
	Leu	Ser	Thr	Lys	Ser	Leu	Val	Ile	Gln	Pro	Arg	Asp	Ile	Leu	Leu	Val
			100					105					110			
20	Ala	Gly	Asn	Pro	Glu	Ile	Leu	Asn	Ala	Val	Tyr	Leu	Gln	Val	Lys	Ser
		115					120					125				
	Asn	Val	Gly	Gln	Phe	Pro	Ala	Pro	Phe	Gly	Lys	Ser	Ile	Tyr	Leu	Tyr
	130					135					140					
	Ile	Asp	Met	Arg	Leu	Gln	Asn	Arg	Lys	Ala	Met	Met	Arg	Asp	Val	Tyr
25	145				150				155					160		
	Gln	Ala	Leu	Phe	Leu	His	Lys	His	Leu	Lys	Ser	Tyr	Lys	Leu	Tyr	Ile
				165					170				175			
	Gln	Val	Leu	His	Pro	Thr	Ser	Pro	Lys	Phe	Tyr	His	Lys	Phe	Leu	Ala
			180					185					190			
30	Leu	Glu	Thr	Glu	Ser	Ile	Glu	Val	Asn	Phe	Asp	Phe	Tyr	Arg	Lys	Ser
		195					200					205				
	Phe	Ile	Gln	Lys	Leu	His	Glu	Asp	His	Gln	Lys	Lys	Met	Gly	Leu	Ile
	210					215					220					
	Val	Val	Gly	Arg	Glu	Leu	Phe	Leu	Ser	Lys	Lys	His	Arg	Lys	Ala	Leu
35	225				230				235					240		
	Tyr	Lys	Thr	Ala	Thr	Pro	Val	Tyr	Lys	Thr	Asn	Thr	Ser	Gly	Leu	Ser
			245					250					255			
	Lys	Thr	Ser	Gln	Ser	Val	Val	Val	Leu	Asn	Glu	Ser	Leu	Asp	Ile	Asn
			260					265					270			
40	Glu	Asp	Met	Ser	Ser	Val	Ile	Phe	Asp	Val	Ser	Met	Gln	Met	Asp	Leu
		275						280					285			
	Gly	Leu	Leu	Leu	Tyr	Asp	Phe	Asp	Pro	Asn	Lys	Arg	Tyr	Lys	Asn	Glu
	290					295						300				
	Ile	Val	Asn	His	Tyr	Glu	Asn	Leu	Ala	Asn	Ala	Phe	Asn	Arg	Lys	Ile
45	305				310				315					320		
	Glu	Ile	Phe	Gln	Thr	Asp	Ile	Arg	Asn	Pro	Ile	Met	Tyr	Leu	Asn	Ser
				325				330					335			

Leu Arg Asn Pro Ile Leu His Phe Met Pro Phe Glu Glu Cys Ile Thr
340 345 350
His Thr Arg Phe Trp Trp Phe Leu Ser Thr Lys Val Glu Lys Leu Ala
355 360 365
5 Phe Leu Asn Asp Asp Asn Pro Gln Ile Phe Ile Pro Val Ala Glu
370 375 380

(2) INFORMATION FOR SEQ ID NO:89:

(i) SEQUENCE CHARACTERISTICS:

10

- (A) LENGTH: 517 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

15

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 51...464
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:89:

20	AGATTTTCATT CGAGGTAGAA AATACATTGA AAAAGCGTGT GAATTAAACG	ATG GTA	56
		Met Val	
		1	
	GGG GGT GGA ACG GTA AAA AAA GAC TTG AAG AAA GCC ATT CAA TAC TAT		104
	Gly Gly Gly Thr Val Lys Lys Asp Leu Lys Lys Ala Ile Gln Tyr Tyr		
	5 10 15		
25	GTT AAA GCG TGT GAA TTG AAT GAA ATG TTT GGG TGT CTG TCA TTA GTT		152
	Val Lys Ala Cys Glu Leu Asn Glu Met Phe Gly Cys Leu Ser Leu Val		
	20 25 30		
30	TCG AAC TCT CAA ATA AAC AAA CAA AAA CTC TTT CAA TAT CTC TCT AAA		200
	Ser Asn Ser Gln Ile Asn Lys Gln Lys Leu Phe Gln Tyr Leu Ser Lys		
	35 40 45 50		
	GCT TGT GAA TTA AAT AGT GGT AAT GGA TGT AGG TTT TTA GGG GAT TTT		248
	Ala Cys Glu Leu Asn Ser Gly Asn Gly Cys Arg Phe Leu Gly Asp Phe		
	55 60 65		
35	TAT GAG AAT GGA AAA TAT GTA AAA AAG GAT TTA AGA AAA GCT GCT CAA		296
	Tyr Glu Asn Gly Lys Tyr Val Lys Lys Asp Leu Arg Lys Ala Ala Gln		
	70 75 80		
	TAC TAC TCT AAA GCT TGT GGA TTA AAT GAT CAA GAT GGG TGT TTA ATA		344
	Tyr Tyr Ser Lys Ala Cys Gly Leu Asn Asp Gln Asp Gly Cys Leu Ile		

85

90

95

CTA GGA TAT AAG CAA TAT GCT GGC AAG GGC GTA GTC AAA AAT GAA AAA 392
 Leu Gly Tyr Lys Gln Tyr Ala Gly Lys Gly Val Val Lys Asn Glu Lys
 100 105 110

5 CAA GCG GTG AAA ACC TTT GAA AAG GCT TGT AGG TTA GGA TCT GAA GAC 440
 Gln Ala Val Lys Thr Phe Glu Lys Ala Cys Arg Leu Gly Ser Glu Asp
 115 120 125 130

GCA TGT GGT ATT TTA AAC AAC TAC TAGATTGAA ATAAATGCTG TTTTITAGCT 494
 Ala Cys Gly Ile Leu Asn Asn Tyr
 10 135

GGCTTTCATG TTTTGTAAAC CCC 517

(2) INFORMATION FOR SEQ ID NO:90:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 138 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:90:

Met Val Gly Gly Gly Thr Val Lys Lys Asp Leu Lys Lys Ala Ile Gln
 1 5 10 15
 Tyr Tyr Val Lys Ala Cys Glu Leu Asn Glu Met Phe Gly Cys Leu Ser
 20 25 30
 25 Leu Val Ser Asn Ser Gln Ile Asn Lys Gln Lys Leu Phe Gln Tyr Leu
 35 40 45
 Ser Lys Ala Cys Glu Leu Asn Ser Gly Asn Gly Cys Arg Phe Leu Gly
 50 55 60
 Asp Phe Tyr Glu Asn Gly Lys Tyr Val Lys Lys Asp Leu Arg Lys Ala
 30 65 70 75 80
 Ala Gln Tyr Tyr Ser Lys Ala Cys Gly Leu Asn Asp Gln Asp Gly Cys
 85 90 95
 Leu Ile Leu Gly Tyr Lys Gln Tyr Ala Gly Lys Gly Val Val Lys Asn
 100 105 110
 35 Glu Lys Gln Ala Val Lys Thr Phe Glu Lys Ala Cys Arg Leu Gly Ser
 115 120 125
 Glu Asp Ala Cys Gly Ile Leu Asn Asn Tyr
 130 135

(2) INFORMATION FOR SEQ ID NO:91:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1663 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
(B) LOCATION: 68...1600
(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:91:

AAATGTTAGA AACCTTACA AAACAAGCTA ATATATTCTA TTCAATTTGC CTCAAGGACA 60
AACAAAC ATG AAA AAA CTT CTT TAT ACC ATA CTC GCG CTT CTT TTA ATC 109
15 Met Lys Lys Leu Leu Tyr Thr Ile Leu Ala Leu Leu Leu Ile
1 5 10
GGC CTT TTA ACA ATC TAT CTC ATC CTT TTT ACA GAA TGG GGG AAT AAG 157
Gly Leu Leu Thr Ile Tyr Leu Ile Leu Phe Thr Glu Trp Gly Asn Lys
15 20 25 30
20 ATC ATC GCT TCG TAT ATA GAG AAA AAA ATC AAC CCG AAC GAG CAC TAC 205
Ile Ile Ala Ser Tyr Ile Glu Lys Lys Ile Asn Pro Asn Glu His Tyr
35 40 45
TTG AGC GTT AAA ACC TTT AAA TTG AGA TTC AAC TCT TTG GAT TTT AAA 253
Leu Ser Val Lys Thr Phe Lys Leu Arg Phe Asn Ser Leu Asp Phe Lys
25 50 55 60
GCT CAA GCC AAC GAT GAT TCC ACG CTC ATT CTT AAG GGG GAT TTT TCA 301
Ala Gln Ala Asn Asp Asp Ser Thr Leu Ile Leu Lys Gly Asp Phe Ser
65 70 75
CTT TTA AAG CAA AGC GTA AAT TTG AAT TAC CAT ATA GAT ATT AAA GAT 349
30 Leu Leu Lys Gln Ser Val Asn Leu Asn Tyr His Ile Asp Ile Lys Asp
80 85 90
TTA CGC TCT TTC AAA GAA TGG ATA CCC TAC CCT TTA AGG GGG GCT GTT 397
Leu Arg Ser Phe Lys Glu Trp Ile Pro Tyr Pro Leu Arg Gly Ala Val
95 100 105 110
35 ATC ACT TCT GGG AAT ATT AAA GGG CAT AGA AAA GCC CTT ATG ATT CAA 445
Ile Thr Ser Gly Asn Ile Lys Gly His Arg Lys Ala Leu Met Ile Gln
115 120 125
GGC GTC TCT AAT GTG GCT CAA TCC CAC ACT GCC TAC AAT GCC CTT TTA 493

	Gly	Val	Ser	Asn	Val	Ala	Gln	Ser	His	Thr	Ala	Tyr	Asn	Ala	Leu	Leu	
				130					135					140			
	GAT	GAT	TTC	AAG	CTT	TCT	CGC	TTA	AAT	TTG	AAC	GCA	CAA	GAC	GCC	AAT	541
5	Asp	Asp	Phe	Lys	Leu	Ser	Arg	Leu	Asn	Leu	Asn	Ala	Gln	Asp	Ala	Asn	
			145					150					155				
	TTA	GAA	GAT	TTG	CTT	TAT	TTA	ATC	AAT	CGC	CCC	GCT	TAT	GCG	AAC	GCA	589
	Leu	Glu	Asp	Leu	Leu	Tyr	Leu	Ile	Asn	Arg	Pro	Ala	Tyr	Ala	Asn	Ala	
		160					165				170						
10	AAA	GTG	TCC	TTA	CAG	GCG	GAT	TTT	AAC	TCT	CTA	AAG	CCT	TTA	GAG	GGG	637
	Lys	Val	Ser	Leu	Gln	Ala	Asp	Phe	Asn	Ser	Leu	Lys	Pro	Leu	Glu	Gly	
	175				180					185				190			
	CAT	TTG	ATC	CTA	ACA	GCT	AAT	AAC	GCT	TTA	ATC	AAT	AAC	GCC	CTA	ATC	685
	His	Leu	Ile	Leu	Thr	Ala	Asn	Asn	Ala	Leu	Ile	Asn	Asn	Ala	Leu	Ile	
				195					200				205				
15	AAT	CAA	ATT	TTT	CAT	TTA	AAC	CTT	AAA	GAC	ACG	CTT	GTT	TTC	AGC	CTC	733
	Asn	Gln	Ile	Phe	His	Leu	Asn	Leu	Lys	Asp	Thr	Leu	Val	Phe	Ser	Leu	
			210					215				220					
20	TCG	CAT	TCA	AGC	GAC	TTT	AAA	GGA	AAC	AAA	GCC	ATC	AGC	GAT	ACC	ACC	781
	Ser	His	Ser	Ser	Asp	Phe	Lys	Gly	Asn	Lys	Ala	Ile	Ser	Asp	Thr	Thr	
			225				230				235						
	CTG	ACT	AGC	CCT	TTA	GCC	AAT	TTC	AAA	GCC	CTA	AAA	AGC	GAA	TAC	CTT	829
	Leu	Thr	Ser	Pro	Leu	Ala	Asn	Phe	Lys	Ala	Leu	Lys	Ser	Glu	Tyr	Leu	
		240				245				250							
25	TTC	TCT	ATT	TTA	AAA	CTC	AAC	GCC	CCC	TAC	ACT	TTA	GAA	ATC	CCC	AAT	877
	Phe	Ser	Ile	Leu	Lys	Leu	Asn	Ala	Pro	Tyr	Thr	Leu	Glu	Ile	Pro	Asn	
	255			260				265				270					
	CTA	GCC	AAA	CTC	TAT	AAC	ATT	ACC	AAC	CAC	CCC	TTA	AAA	GGG	AGC	TTG	925
	Leu	Ala	Lys	Leu	Tyr	Asn	Ile	Thr	Asn	His	Pro	Leu	Lys	Gly	Ser	Leu	
			275					280				285					
30	ACT	TTA	AAA	GGC	GCT	ATA	GAA	CAA	AGC	CCC	AAA	CTT	TTA	AAA	GTC	AGC	973
	Thr	Leu	Lys	Gly	Ala	Ile	Glu	Gln	Ser	Pro	Lys	Leu	Leu	Lys	Val	Ser	
			290					295				300					
	GGC	CAT	TCA	AAT	TTA	CTA	GAC	GGC	GCG	CTG	GAT	TTC	ACG	CTT	TTA	AAT	1021
	Gly	His	Ser	Asn	Leu	Leu	Asp	Gly	Ala	Leu	Asp	Phe	Thr	Leu	Leu	Asn	
35			305				310				315						
	AAA	GAT	TTG	AAA	GGG	CGT	TTT	TCC	AAT	ATT	TCC	ACT	TTA	AAA	GCT	TTA	1069
	Lys	Asp	Leu	Lys	Gly	Arg	Phe	Ser	Asn	Ile	Ser	Thr	Leu	Lys	Ala	Leu	
		320				325					330						

	GAT TTA TTC CAT TAC CCT AAG TTT TTC CAA TCC GTT GCA GAC GCT AAT	1117
	Asp Leu Phe His Tyr Pro Lys Phe Phe Gln Ser Val Ala Asp Ala Asn	
	335 340 345 350	
	TTG GAT TAT GAT CTT ATC GCT AAG CAA GGC GTA TTG AAA GCC CGC CTA	1165
5	Leu Asp Tyr Asp Leu Ile Ala Lys Gln Gly Val Leu Lys Ala Arg Leu	
	355 360 365	
	AAA AAC GCA AGA TTC CTC AAA AAT GCA TTC AGC GAT TTT CTC TAC TCC	1213
	Lys Asn Ala Arg Phe Leu Lys Asn Ala Phe Ser Asp Phe Leu Tyr Ser	
	370 375 380	
10	ATT TCT AAA TTT GAT ATT ACA AAA GAA ATT TAT AAC GAT GCC AAT CTG	1261
	Ile Ser Lys Phe Asp Ile Thr Lys Glu Ile Tyr Asn Asp Ala Asn Leu	
	385 390 395	
	GTA AGC CAA ATC AAC CAG CAA CGC CTG CTC TCT GAT CTG AGT TTA AAA	1309
	Val Ser Gln Ile Asn Gln Gln Arg Leu Leu Ser Asp Leu Ser Leu Lys	
15	400 405 410	
	AGC CCC AAA ACC CAA TTG AAA ATC CAT AAC GGT TTG TTG GAT TTA AAC	1357
	Ser Pro Lys Thr Gln Leu Lys Ile His Asn Gly Leu Leu Asp Leu Asn	
	415 420 425 430	
	ACC AAA CAA ATG AAC ATG CTC ATG GAT GCG GAA ATT TTA AAA TTC ATT	1405
20	Thr Lys Gln Met Asn Met Leu Met Asp Ala Glu Ile Leu Lys Phe Ile	
	435 440 445	
	TTT AAA ATG AAA CTT CAA GGC AAC ATG CAC CAG CCA AAA TTT TCT CTC	1453
	Phe Lys Met Lys Leu Gln Gly Asn Met His Gln Pro Lys Phe Ser Leu	
	450 455 460	
25	ATT TTA AAC GAA AAA GCC ATT CAG CAA AAC TTG CAA CAA GGC TTG AAA	1501
	Ile Leu Asn Glu Lys Ala Ile Gln Gln Asn Leu Gln Gln Gly Leu Lys	
	465 470 475	
	GAA ATC TTA AAA AAC GAC ACC CTT AAA AAA GGT TTA GAT CAT TTG CTT	1549
	Glu Ile Leu Lys Asn Asp Thr Leu Lys Lys Gly Leu Asp His Leu Leu	
30	480 485 490	
	AAA GAT GAT AAG CTC AAA GAA AAG CTT GAA AAA GGG CTT AAG GGG CTT	1597
	Lys Asp Asp Lys Leu Lys Glu Lys Leu Glu Lys Gly Leu Lys Gly Leu	
	495 500 505 510	
	TTT TAAAAATTTT AAAGGATAGA AATGGCGCAC ATTTTAGTTA GCGGGGCGAC TTCAGG	1656
35	Phe	
	GTTTGGGA	1663

(2) INFORMATION FOR SEQ ID NO:92:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 511 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:92:

10 Met Lys Lys Leu Leu Tyr Thr Ile Leu Ala Leu Leu Leu Ile Gly Leu
1 5 10 15
Leu Thr Ile Tyr Leu Ile Leu Phe Thr Glu Trp Gly Asn Lys Ile Ile
20 25 30
Ala Ser Tyr Ile Glu Lys Lys Ile Asn Pro Asn Glu His Tyr Leu Ser
15 35 40 45
Val Lys Thr Phe Lys Leu Arg Phe Asn Ser Leu Asp Phe Lys Ala Gln
50 55 60
Ala Asn Asp Asp Ser Thr Leu Ile Leu Lys Gly Asp Phe Ser Leu Leu
65 70 75 80
20 Lys Gln Ser Val Asn Leu Asn Tyr His Ile Asp Ile Lys Asp Leu Arg
85 90 95
Ser Phe Lys Glu Trp Ile Pro Tyr Pro Leu Arg Gly Ala Val Ile Thr
100 105 110
Ser Gly Asn Ile Lys Gly His Arg Lys Ala Leu Met Ile Gln Gly Val
115 120 125
25 Ser Asn Val Ala Gln Ser His Thr Ala Tyr Asn Ala Leu Leu Asp Asp
130 135 140
Phe Lys Leu Ser Arg Leu Asn Leu Asn Ala Gln Asp Ala Asn Leu Glu
145 150 155 160
30 Asp Leu Leu Tyr Leu Ile Asn Arg Pro Ala Tyr Ala Asn Ala Lys Val
165 170 175
Ser Leu Gln Ala Asp Phe Asn Ser Leu Lys Pro Leu Glu Gly His Leu
180 185 190
Ile Leu Thr Ala Asn Asn Ala Leu Ile Asn Asn Ala Leu Ile Asn Gln
195 200 205
35 Ile Phe His Leu Asn Leu Lys Asp Thr Leu Val Phe Ser Leu Ser His
210 215 220
Ser Ser Asp Phe Lys Gly Asn Lys Ala Ile Ser Asp Thr Thr Leu Thr
225 230 235 240
40 Ser Pro Leu Ala Asn Phe Lys Ala Leu Lys Ser Glu Tyr Leu Phe Ser
245 250 255
Ile Leu Lys Leu Asn Ala Pro Tyr Thr Leu Glu Ile Pro Asn Leu Ala
260 265 270
Lys Leu Tyr Asn Ile Thr Asn His Pro Leu Lys Gly Ser Leu Thr Leu
275 280 285
45 Lys Gly Ala Ile Glu Gln Ser Pro Lys Leu Leu Lys Val Ser Gly His

	290		295		300	
	Ser Asn Leu Leu Asp Gly	Ala Leu Asp Phe Thr	Leu Leu Asn Lys Asp			
	305	310	315	320		
	Leu Lys Gly Arg Phe Ser	Asn Ile Ser Thr	Leu Lys Ala Leu Asp	Leu		
5		325	330	335		
	Phe His Tyr Pro Lys Phe	Phe Gln Ser Val	Ala Asp Ala Asn	Leu Asp		
		340	345	350		
	Tyr Asp Leu Ile Ala Lys	Gln Gly Val Leu	Lys Ala Arg Leu	Lys Asn		
		355	360	365		
10	Ala Arg Phe Leu Lys Asn	Ala Phe Ser Asp	Phe Leu Tyr Ser	Ile Ser		
		370	375	380		
	Lys Phe Asp Ile Thr Lys	Glu Ile Tyr Asn	Asp Ala Asn Leu	Val Ser		
		385	390	395	400	
	Gln Ile Asn Gln Gln Arg	Leu Leu Ser Asp	Leu Ser Leu Lys	Ser Pro		
15		405	410	415		
	Lys Thr Gln Leu Lys Ile	His Asn Gly Leu	Leu Asp Leu Asn	Thr Lys		
		420	425	430		
	Gln Met Asn Met Leu Met	Asp Ala Glu Ile	Leu Lys Phe Ile	Phe Lys		
		435	440	445		
20	Met Lys Leu Gln Gly Asn	Met His Gln Pro	Lys Phe Ser Leu	Ile Leu		
		450	455	460		
	Asn Glu Lys Ala Ile Gln	Gln Asn Leu Gln	Gln Gly Leu Lys	Glu Ile		
		465	470	475	480	
	Leu Lys Asn Asp Thr Leu	Lys Lys Gly Leu	Asp His Leu Leu	Lys Asp		
25		485	490	495		
	Asp Lys Leu Lys Glu Lys	Leu Glu Lys Gly	Leu Lys Gly Leu	Phe		
		500	505	510		

(2) INFORMATION FOR SEQ ID NO:93:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 947 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
 (B) LOCATION: 292...645
 (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:93:

40	AGTGCATAAA CGCACAGACC CCAAAAATGA AAGCTATTTT TGGCTAGGGC TACACCCTTT	60
	AGAATGGCAA AAGCGCGAAA ATGAAGACAG ACTCTCTGAT TTTGACGCTA TTGCTTCAAA	120
	CCATGCCTCT ATCAGCCTT TAAATTTAGA CTTAACCAGT TATGATGATT TGAAAAGTTT	180
	GGAATCTTGG CATGAGGGAA TGTTAAAGTG AGTAAAAAGC ACCGCTTGGC TTTTITTAGGG	240
	CTAATTGTTG GGGTTCTATT CTTCTTTAGT GCGTGTGAGC ACCGCCTGCA C ATG GGG	297

Met Gly
1

5	TAT TAT TCA GAA GTT ACA GGG GAT TAT TTG TTC AAT TAT AAT TCC ACT	345
	Tyr Tyr Ser Glu Val Thr Gly Asp Tyr Leu Phe Asn Tyr Asn Ser Thr	
	5 10 15	
	ATC GTG GTG GCT TAT GAC AGA AGC GAT GCG ATG ACT TCT TAT TAT ATC	393
	Ile Val Val Ala Tyr Asp Arg Ser Asp Ala Met Thr Ser Tyr Tyr Ile	
	20 25 30	
10	AAT GTG ATT GTT TAT GAA TTG CAA AAA TTA GGC TTT TAC AAT GTC TTC	441
	Asn Val Ile Val Tyr Glu Leu Gln Lys Leu Gly Phe Tyr Asn Val Phe	
	35 40 45 50	
	ACG CAA GCG GAA TTC CCA CTA GAT AAA GCC AAA AAT GTG ATC TAT GCG	489
	Thr Gln Ala Glu Phe Pro Leu Asp Lys Ala Lys Asn Val Ile Tyr Ala	
	55 60 65	
15	CGC ATT GTC CGT AAC ATC TCA GCT GTG CCG TTC TAC CAA TAC AAT TAC	537
	Arg Ile Val Arg Asn Ile Ser Ala Val Pro Phe Tyr Gln Tyr Asn Tyr	
	70 75 80	
	CAA CTG ATT GAT CAA GTC AAT AAG CCT TGT TAT TTT CTT GGG GGG CAG	585
20	Gln Leu Ile Asp Gln Val Asn Lys Pro Cys Tyr Phe Leu Gly Gly Gln	
	85 90 95	
	TTT TAT TGC TCT CAA ACC CTA CGG ATT ATT ACG CTA TCA ATG GCT TTA	633
	Phe Tyr Cys Ser Gln Thr Leu Arg Ile Ile Thr Leu Ser Met Ala Leu	
	100 105 110	
25	GCG AGC AAA TTT TAATGAGTGC TAATTCGCAT TTTATTTTAG ATTGGTATGA TGTGG	690
	Ala Ser Lys Phe	
	115	
	TGTTGCAAAA ACGGGTTTTA TATGTGGATG GGAGCGTGAG CGGGAGGACT TGCGGCTATC	750
	AGATGCTGTA TAGGGATTTG ATTAAAAGCA CGATCAAACG CATTGATTTT AACCGCCCTG	810
	AACGCTACTA CTACAATTTA AGACTGCCCC TTTATCAGCC ATGTTATAGG CAATGAAATG	870
30	GTTATCAGGC GATTGTATCA ATTTTGCGCT AGCCATGTGG TGC GCAATTG CTCTTCTTTA	930
	AAATGCGCTC AAAATAT	947

(2) INFORMATION FOR SEQ ID NO:94:

(i) SEQUENCE CHARACTERISTICS:

- 35 (A) LENGTH: 118 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:94:

```
Met Gly Tyr Tyr Ser Glu Val Thr Gly Asp Tyr Leu Phe Asn Tyr Asn
  1      5      10      15
5 Ser Thr Ile Val Val Ala Tyr Asp Arg Ser Asp Ala Met Thr Ser Tyr
  20      25      30
Tyr Ile Asn Val Ile Val Tyr Glu Leu Gln Lys Leu Gly Phe Tyr Asn
  35      40      45
10 Val Phe Thr Gln Ala Glu Phe Pro Leu Asp Lys Ala Lys Asn Val Ile
  50      55      60
Tyr Ala Arg Ile Val Arg Asn Ile Ser Ala Val Pro Phe Tyr Gln Tyr
  65      70      75      80
Asn Tyr Gln Leu Ile Asp Gln Val Asn Lys Pro Cys Tyr Phe Leu Gly
  85      90      95
15 Gly Gln Phe Tyr Cys Ser Gln Thr Leu Arg Ile Ile Thr Leu Ser Met
  100      105      110
Ala Leu Ala Ser Lys Phe
  115
```

(2) INFORMATION FOR SEQ ID NO:95:

20 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 875 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

25 (ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 348...716
- (D) OTHER INFORMATION:

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:95:

```
TGCGGAGGGA ATGTCTATGA TAAATCTCA GAAAAATTG TAGAAAAAGT GGATAACGGG      60
TTTGAAAAAT TTAAATCCTT TTTTATCT GTTTAAACGC ATTGTTTCGCC CTAGATTCAA      120
ACGCACTTAA AGCAGAGATT AAAGAAGTTT ACCTTAAAGA ATACAAAGAC TTAAAATTAG      180
AAATTGAAAC CATTAACTTA GAAATCCCAG AGCGCTTTTC TAACGCTTCC ATTTTAAGCT      240
35 ATGAATTAAA CGCTTCCAAT AAGCTTAAAA AAGATGGGGT CGTGTTTTTA AGGTTGGAAA      300
ATGATCCTAA TTTACGCTA CCGGTGCGTT ATAGCGTGAT AGGCAGC ATG CAG GCT      356
                                     Met Gln Ala
                                     1

TTT AAA AGC GTT AGC GCG ATT AAA AAA GAT GAA AAC ATC ACC GCT AAT      404
40 Phe Lys Ser Val Ser Ala Ile Lys Lys Asp Glu Asn Ile Thr Ala Asn
   5              10              15
```

AAC ACT CAA AAA GAG CGC ATT TTG TTT GGT GCG CTT TCT AAC CCC TTA 452
Asn Thr Gln Lys Glu Arg Ile Leu Phe Gly Ala Leu Ser Asn Pro Leu
20 25 30 35

	TTA	GAG	GGC	GCG	ATT	GAT	AAA	GTG	AGC	GCG	AAA	AAT	TTT	ATC	CCC	CCT	500
5	Leu	Glu	Gly	Ala	Ile	Asp	Lys	Val	Ser	Ala	Lys	Asn	Phe	Ile	Pro	Pro	
					40					45					50		

AAC ACG CTT TTA AGC ACG GAT AAA ACC CAA GCT TTA ATT ATC GTG CGT 548
Asn Thr Leu Leu Ser Thr Asp Lys Thr Gln Ala Leu Ile Ile Val Arg
55 60 65

10 AAA AAT GAC ATT ATC ACC GGG GTG TAT GAA GAG GGG CAA ATC AGC ATA 596
Lys Asn Asp Ile Ile Thr Gly Val Tyr Glu Glu Gly Gln Ile Ser Ile
70 75 80

GAA ATA AGC CTA AAA GCC CTA GAA AAT GGC GCG CTT AAT CAA ATC ATT 644
Glu Ile Ser Leu Lys Ala Leu Glu Asn Gly Ala Leu Asn Gln Ile Ile

15 85 90 95

CAA GCG AAA AAT TTA GAA AGC AAT AAA ATA CTC AAA GCA AAA GTG TTG 692
Gln Ala Lys Asn Leu Glu Ser Asn Lys Ile Leu Lys Ala Lys Val Leu
100 105 110 115

20 AGC AGC TCT AAA GCG CAA ATC TTA TAAAGGACAT TCATGAAATT GGTTTTAGGC 746
Ser Ser Ser Lys Ala Gln Ile Leu
120

ATCAGTGGAG	CGAGCGGGAT	ACCCCTAGCC	TTGCGGTTTT	TAGAAAAATT	ACCCAAAGAA	806
ATTGAAGTTT	TTGTCGTGGC	GTCTAAAAAC	GCGCATGTCG	TGGCGTTAGA	AGAATCTAAT	866
ATTAACCTT						875

25 (2) INFORMATION FOR SEQ ID NO:96:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 123 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:96:

35 Met Gln Ala Phe Lys Ser Val Ser Ala Ile Lys Lys Asp Glu Asn Ile
1 5 10 15
Thr Ala Asn Asn Thr Gln Lys Glu Arg Ile Leu Phe Gly Ala Leu Ser
20 25 30
Asn Pro Leu Leu Glu Gly Ala Ile Asp Lys Val Ser Ala Lys Asn Phe

	35		40		45											
	Ile	Pro	Asn	Thr	Leu	Leu	Ser	Thr	Asp	Lys	Thr	Gln	Ala	Leu	Ile	
	50		55		60											
	Ile	Val	Arg	Lys	Asn	Asp	Ile	Ile	Thr	Gly	Val	Tyr	Glu	Glu	Gly	Gln
5	65		70		75		80									
	Ile	Ser	Ile	Glu	Ile	Ser	Leu	Lys	Ala	Leu	Glu	Asn	Gly	Ala	Leu	Asn
		85		90		95										
	Gln	Ile	Ile	Gln	Ala	Lys	Asn	Leu	Glu	Ser	Asn	Lys	Ile	Leu	Lys	Ala
		100		105		110										
10	Lys	Val	Leu	Ser	Ser	Ser	Lys	Ala	Gln	Ile	Leu					
	115		120													

(2) INFORMATION FOR SEQ ID NO:97:

(i) SEQUENCE CHARACTERISTICS:

- 15 (A) LENGTH: 394 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:

- 20 (A) NAME/KEY: Coding Sequence
(B) LOCATION: 160...345
(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:97:

	GGCATCACTT	TTAACATGAC	CCCTTCTCCA	GGCGCGACGA	GTTGTTTGCA	AAACGCCCTT	60
25	GTGGATTCCC	AAGAAATCGC	TGCGTATTTG	GGCGAGAGCT	TTGAATTAGA	ACGCTTTTAT	120
	AAAGATTTAT	CCCCAGAAGA	ATTGGAAAAT	TAAAAACGC	ATG CAA AAA	GAA CAA	174
				Met Gln Lys	Glu Gln		
				1	5		
	GAA GCC CAA	GAA ATC GCT	AAA AAA GCC	GTT AAA ATC	GTG TTT TTT	TTA	222
30	Glu Ala Gln	Glu Ile Ala	Lys Lys Ala	Val Lys Ile	Val Phe Phe	Leu	
		10		15	20		
	GGG CTT GTG	GTG GTG CTT	TTG ATG ATG	ATA AAC CTT	TAC ATG CTC	ATC	270
	Gly Leu Val	Val Val Leu	Leu Met Met	Ile Asn Leu	Tyr Met Leu	Ile	
		25		30	35		
35	AAT CAA ATC	AAC GCG AGC	GCT CAA ATG	AGC CAC CAA	ATC AAA AAG	ATA	318
	Asn Gln Ile	Asn Ala Ser	Ala Gln Met	Ser His Gln	Ile Lys Lys	Ile	
		40		45	50		
	GAA GAA AGG	CTT AAT CAG	GAG CAA AAA	TAAAAAAGGC	TTTTTGGTAT	TTTTACG	372
40	Glu Glu Arg	Leu Asn Gln	Glu Gln Lys				
	55		60				

ATCAAATAGT AAAGAGCTTA TC

394

(2) INFORMATION FOR SEQ ID NO:98:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 62 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:98:

Met Gln Lys Glu Gln Glu Ala Gln Glu Ile Ala Lys Lys Ala Val Lys
1 5 10 15
Ile Val Phe Phe Leu Gly Leu Val Val Val Leu Leu Met Met Ile Asn
20 25 30
Leu Tyr Met Leu Ile Asn Gln Ile Asn Ala Ser Ala Gln Met Ser His
35 40 45
Gln Ile Lys Lys Ile Glu Glu Arg Leu Asn Gln Glu Lys
50 55 60

(2) INFORMATION FOR SEQ ID NO:99:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 516 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:99:

GTTTGATTTA GTTCAAGAGC TTTTAGAAGA ATTTTGTCAA AGCGGGGCTA AAGAGATTTT 60
AGAAAAGGCG CAGTTGTTTT AATGCGTTTG TTTATCGCGC TAGTTTTGTT TTGGTGGTGG 120
TTAAGCTTGA ACGCTAAAGA AGCGGATTTT ATCTCTGATT TAGAATACGG GATGGCTCTT 180
TATAAAAACC CTAGGGGTGT TCGGTGCGCG AAATGCCATG GCATTAAAGG CGAACAACAA 240
GAAATCACCT TTTATTATGA AAAAGGCGAG AAAAAAATCC TCTACGCCCC TAAAAATCAAC 300
CATTTGGATT TTAAACCTT TAAAGACGCC TTGAGTTTAG GCAAAGGCAT GATGCCTAAA 360
TACAATCTCA ATTTAGAAGA AATCCAAGCG ATTTATCTTT ATATCATCTC TTTAGAGCAT 420
AAAGAAGAGC GTAAGGATTC TCCTAAGCCT TAATCAAAGC GCTTGATTTA TGCTAAAATG 480
GAGCGTTGCA TTTTGTGTTT GATTAAAGAA GGGTTC 516

(2) INFORMATION FOR SEQ ID NO:100:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 709 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:100:

```
5  TAAGGGATAT TGCTAACGAT TAAGCTGTAT TGGAAGAGTT TATTTTGCAA GAATTAATCT 60
   TGCCTTGTGT GATTAGTAAC ACAAGGCAAG TGTGATAAAC CCTACTACAA TTTCAATTCA 120
   AGGAGCCTAA CTAAAATAAA ATGAACAATT TCAGTTAGGG CTTTATTATA GCAAAAATTA 180
   TCTAAGATTA CAAAGGGTAG CGTTTCTGTT TTTGGATTTA GAGCGTTATT TTGATTGTTT 240
   TGAGTTTAAT TTACTTTTTG TTTAATAATA AATCTTAACT ATCATAAATG TACAATTAAA 300
10 GTATTTAAAA AAATTTTAAA ACAAAGGAT ATAAAATGAA AACCATTAGA AATAGCGTGT 360
   TTATTGGAGC GTCTTTACTC GGCGGTTGCG CTAGCGTTGA GGCTTATTTT GACGCTTTGC 420
   ATGTTGCTCG CGTTAAAGAC GCTTGTTTAT AGAAAAAGAA GCACACCACA CGCCCAAAGA 480
   CTTTGATAGC CCTTACCACA CTGACTAAAC CGGCACTAGG TTTTAGTTGG GGGTTTTTAG 540
   GGGTGTTATT TTAGATACTC TCTGTTCCCT TAAAGAAAAT AAATTTCTAC CATAAAATAA 600
15 AATCTTAAAT TAAGGCGACT AAAACCCAC TTTTAAAAAA TTAAAAAGCG TTAAGTAAGA 660
   CTTATCCAAA AAGCAAAGAA AATCAATTTT TCCAACCACT TTTTTTAAG 709
```

(2) INFORMATION FOR SEQ ID NO:101:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:101:

```
CGCGGATCCG AAATAGGGTT GTTTTAAATT TTC 33
```

(2) INFORMATION FOR SEQ ID NO:102:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:102:

```
CCGCTCGAGT TAAAAAAGA GTTTGTATAA 30
```

(2) INFORMATION FOR SEQ ID NO:103:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 25 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:103:

5 GGGGATCCTT GGTAGAATTG AATCA

25

(2) INFORMATION FOR SEQ ID NO:104:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:104:

GGAATTCCTA AAACAAGAAC GCG

23

(2) INFORMATION FOR SEQ ID NO:105:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:105:

GGGGATCCTT TTTTCAAAAA CAATA

25

(2) INFORMATION FOR SEQ ID NO:106:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:106:

GGAATTCTCA CATTGTTTTG CTC

23

30

(2) INFORMATION FOR SEQ ID NO:107:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:107:

GCGGATCCCA ATTTCAAAAA GCC

23

(2) INFORMATION FOR SEQ ID NO:108:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:108:

15 CCGCTCGAAC TAAAACTAT AAACG

25

(2) INFORMATION FOR SEQ ID NO:109:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:109:

CGCGGATCCG AGATTTTGAA AGGTTGGTAA TG

32

(2) INFORMATION FOR SEQ ID NO:110:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:110:

CCGCTCGAGC TACATCCTTT TACTATAACC

30

(2) INFORMATION FOR SEQ ID NO:111:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:111:

GCGGATCCGG GTATTATTCA GAAG

24

(2) INFORMATION FOR SEQ ID NO:112:

10

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

15

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:112:

CCGCTCGAGT TAAAATTTGC TCGC

24